Dictionaries of Medieval Germanic Languages

International Medieval Research

Selected Proceedings of the International Medieval Congress
University of Leeds

Academic Editorial Board
Simon Forde, Joyce Hill, Luc Jocqué

Production Manager
Amanda Banton

International Medieval Research

Dictionaries of Medieval Germanic Languages

A Survey of Current Lexicographical Projects

Selected Proceedings of the
International Medieval Congress
University of Leeds,
4-7 July 1994

edited by
K.H. van Dalen-Oskam, K.A.C. Depuydt,
W.J.J. Pijnenburg and T.H. Schoonheim

Turnhout, Brepols
1997

Articles appearing in this volume are indexed in
"International Medieval Bibliography"

© 1997 – BREPOLS
Printed in Belgium
D/1997/0095/42
ISBN 2-503-50601-1
All rights reserved. No part of this publication may be reproduced,
stored in a retrieval system
or transmitted, in any form or by any means, electronic, mechanical, photocopying,
recording, or otherwise, without the prior permission of the publisher.

Contents

Editors' Introduction vii

I Forschungsberichte

Dutch

Vroegmiddelnederlands Woordenboek. 1200-1300 (Dictionary of Early Middle Dutch: 1200-1300)
Willy Pijnenburg 3

CD-ROM Middelnederlands (Dictionary of Middle Dutch)
Dirk Geirnaert 11

Historisch Woordenboek der Nederlandse Rechtstaal. 14de-18de eeuw (Historical Dictionary of Dutch Legal Terminology: 14th-18th Century)
Christel Verhas 25

English

A Thesaurus of Old English
Jane Roberts and Christian Kay 31

A Thesaurus of Middle English
Jane Roberts and Louise Sylvester 41

Historical Thesaurus of English
Christian Kay and Irené Wotherspoon 47

Dictionary of Old English
Antonette diPaolo Healey 55

Middle English Dictionary
Robert E. Lewis 63

German

Mittelhochdeutsches Wörterbuch. 1050-1350 (Middle High German Dictionary: 1050-1350)
Kurt Gärtner and Ralph Plate 75

Norse

Ordbog over det Norrøne Prosasprog (A Dictionary of Old Norse Prose)
Christopher Sanders 81

Gammalnorsk Ordboksverk (Dictionary of Old Norwegian c. 1200-1550)
Bjørn Eithun 89

Ordbok over Skriftspråket I Norge 1350-1550 (Dictionary of the Written Language in Norway 1350-1550)
Erik Simensen 99

II Papers

On the Preparation of a Dictionary of Middle Norwegian
Erik Simensen 107
The Vroegmiddelnederlands Woordenboek *and the Computer*
Tanneke Schoonheim 117
The Importance of Legal History for Old Frisian Lexicography
Oebele Vries 135
Encyclopedic Aspects of the Historical Dictionary and the Question of Context
Jeffrey L. Singman 149
Identification of the Veins and Arteries in Middle English Medical Literature
James M. Girsch 159
The Lexicographer's Dilemma: Handling Mistranslations, Misconstruings and Paranomasia in Source-Texts
Elizabeth S. Girsch 177
Lexicography and Philology
Karina van Dalen-Oskam and Katrien Depuydt 189

Editors' Introduction

In 1994 the University of Leeds hosted the first *International Medieval Congress* (4-7 July). Three sessions were dedicated to "Dictionaries of Medieval Languages". They dealt with the following themes: "Projects", in which several dictionary projects were presented, "Historical Background", focusing on the importance of extensive knowledge of the historical background when working on a dictionary of a medieval language and "Scribes and Scholars", elaborating on the philological problems a lexicographer encounters when working with old handwritten source material and its editions.

The main purpose of these sessions was to offer participants an opportunity to get acquainted with other lexicographical projects. The papers and the discussions that followed soon led to further debate on how to handle how certain lexicographical and especially computer-related problems.

As organisers of these three sessions at IMC'94, we concluded that a publication of the presented papers would not only be very useful, but that it would also benefit from a complete survey of the current lexicographical projects on Medieval Germanic Languages. We then contacted the editors of the different projects known to us, urging them to write a *Forschungsbericht* for this volume. We are very happy that so many projects have been willing to participate.

The first part of this volume contains a description of three projects on the Middle Dutch language (one of which also covers Early Modern Dutch), five on Old or Middle English, one on Middle High German[1] and three on Old Norse or Old Norwegian. Each *Forschungsbericht* has the following structure: information about the dictionary (title, editors, institute, address), contents of the dictionary (type, subject/corpus, described period, described area, example of an entry), history of the project and planning (short history of the project, (planned) year of publication, form of publication), lexicographical tools and methods (the hardware, the soft-

[1]For *Forschungsberichte* of other German projects we refer to *Deutschsprachige Wörterbuchprojekte an Akademien und Universitäten* (Göttingen, 1996).

ware), special remarks. The second part presents the Leeds papers, in the order of the above-mentioned session themes.

We believe that this volume contains valuable information, not only for editors of existing lexicographical projects but also for editors of future projects. We expect that it will also give non-lexicographers a better insight into modern historical lexicography.

I Forschungsberichte

Vroegmiddelnederlands Woordenboek 1200-1300

Dictionary of Early Middle Dutch
1200-1300

Information about the Dictionary

EDITORS

Editor-in-Chief: W.J.J. Pijnenburg
Editors: K.A.C. Depuydt
K.H. van Dalen-Oskam
T.H. Schoonheim

INSTITUTE

Instituut voor Nederlandse Lexicologie

ADDRESS

Correspondence is to be sent to the secretary:
Th.P.F. Wortel
Instituut voor Nederlandse Lexicologie
Afdeling Vroegmiddelnederlands Woordenboek
Postbus/PO Box 9515
NL 2300 RA LEIDEN
The Netherlands

e-mail: Dick@rulxho.leidenuniv.nl
tel.: +31 (71) 527-2476

Contents of the Dictionary

TYPE

The *Vroegmiddelnederlands Woordenboek* (*VMNW*) is an alphabetically ordered, scientific, diachronic period-dictionary. The c. 28,000 entries give information on grammatical status, number of occurrences per genre, oldest attestation, etymology, inflection and meaning. The latter is illustrated with quotations. The spelling of the entries is normalised. Nouns and pronouns are given in the nominative singular, verbs in the infinitive.

SUBJECT/CORPUS

The dictionary is based on a closed corpus of early Middle Dutch texts, written between 1200 and 1301 and passed down in thirteenth-century manuscripts. The greater part dates from the last quarter of the thirteenth century. Approximately 1,650,000 tokens will eventually be described in c. 28,000 entries. The majority of the texts have been edited diplomatically by M. Gysseling. They were published between 1977 and 1987 in the two series of the *Corpus van Middelnederlandse teksten (tot en met het jaar 1300)* (Corpus of Early Middle Dutch Texts up to and including the year 1300), with word indexes to the texts. The first series contains nine volumes with official documents, such as charters, guidelines, testaments, etc. and the second series six volumes with literary texts and a few *artes* texts. To Gysseling's material a couple of already published texts have been added, like the *Glossarium Bernense*, an early thirteenth-century glossary and the *Diatessaron Leodiense*.

Every single occurrence in the corpus is being examined and described, including personal names and latinised forms of originally Dutch words (see below). In the eventual dictionary article, of course, not every occurrence is quoted.

DESCRIBED PERIOD

The period the dictionary describes begins with the year 1200 AD, from which, according to scholars, date the fragments we have of the Middle Dutch version of Henric van Veldeke's *Sente Servas*. The youngest (official) documents date from the 31 December 1300. Unlike official documents, which usually have a very exact date (day,

month and year) the literary texts can only be dated approximately, sometimes only with a terminus post quem and a terminus ante quem.

DESCRIBED AREA

The thirteenth-century Dutch-speaking area differs slightly from the present-day Dutch-speaking area, which is the Netherlands and Flanders (Belgium). Medieval Dutch was spoken in the Dutch provinces of North- and South-Holland, Utrecht, Sealand, parts of Guelderland, the Lower Rhine area (Cleves and surroundings; partly belonging to Germany now), Limburg (now divided between the Netherlands and Belgium; without the extreme south-eastern part where the language was Middle High German), Brabant and Flanders (now divided between the Netherlands, Belgium and France). In the Lower Rhine Dutch was spoken until the middle of the last century, and in some parts of French Flanders (e.g. Dunkirk) until the first half of this century.

Excluded are the Frisian-speaking areas in the Netherlands (now the province of Friesland) and the Middle Low German-speaking areas in the Netherlands (now Groningen, Drente, Overijsel, parts of Guelderland).

EXAMPLE OF AN ENTRY

Each entry in the *VMNW* contains the following information: first the headword, in normalised spelling, followed by the word class. Next come a summary of the various meanings of the headword and the total number of occurrences of the headword in the corpus. The information in the summary should be sufficient for a user who only wants to have a quick look at the meaning of a word.

The dictionary article continues with an indication whether or not the entry occurs in the *Middelnederlandsch Woordenboek* (Dictionary of Middle Dutch, c. 1150-1550; 11 volumes), and subsequently mentions the number of occurrences per genre, i.e. lexicographical material, official documents or literary texts, followed by provenance and date of the document containing the oldest occurrence. Spelling variation, sometimes with specification of the number of occurrences of the major variants and – if relevant – remarks on morphology or etymology and information on the geographical distribution of the headword come next.

Finally the semantic and/or syntactic analysis of the headword is given, illustrated with several examples, including the oldest attestation.

History of the Project and Planning

SHORT HISTORY OF THE PROJECT

The *VMNW* is in fact the follow-up of the editorial project of the above-mentioned *Corpus van Middelnederlandse teksten*. A dictionary was believed to be the best tool to unlock this *Corpus* for further scientific research. In 1988, the Board of the Institute for Dutch Lexicology (in which both the Belgian and Dutch government participate) assigned Dr Pijnenburg, who also participated in the *Corpus* project, to the dictionary project as editor-in-chief. He was given three editors, one corrector and ten years' time to complete the work. Dr Pijnenburg then made the design for the dictionary, taking into account the desires future users were enabled to express through a preliminary poll.

Every token of the corpus was then tagged in order to function optimally in a database (described extensively in the article of T.H. Schoonheim, elsewhere in this book). An Advisory Board of thirteen specialists in the field of Middle Dutch language and literature, onomastics and diachronic linguistics was added to advise the editorial team. The project started on 1 January 1989.

(PLANNED) YEAR OF PUBLICATION

The compilation of the dictionary is scheduled to be finished on 31 December 1998. 1999 will be used for correction and for adaptation of the spelling to the new official guidelines, issued in 1996. The dictionary is expected to be available in the concluding year of this millennium, 2000.

FORM OF PUBLICATION

The *VMNW* will be published both in print and on CD-ROM. The printed version will consist of three or four volumes of about 700 pages each. A detailed *Introduction* will give a short history of the project, an explanation of the different information categories in each

entry, their position and their purpose, and also elaborate on the spelling system created for the entries so as to understand their alphabetical ordering. Of course, there will also be added a detailed list of all sources used and of the literature cited in the *VMNW*.

In the meantime, it is already possible to consult the dictionary in progress, since finished articles have already been printed in preliminary issues (some 22,000 in total on 31 December 1996). Scholars who are interested in doing so are requested to contact the secretary of the *VMNW*. It is not unimportant to mention that the editors have chosen not to work in strict alphabetical order, but more or less thematically, whenever possible.

Lexicographical Tools and Methods

HARDWARE

Until 1997 the editors worked with VAX VS 3100 work stations with a 21" colour screen, both from Digital Equipment Company. The bigger screen makes it possible to work with several windows at the same time, thus enabling the editors to manipulate simultaneously the separate parts of the computer system used. On 1 January 1997, they were replaced by personal computers (Targa TO P200) with a 2.1 GB Harddisk, 32 MB main memory and 2 MB VRAM and a 21" colour monitor (NEC Multisync XP21). Also included is an 8-speed CD-ROM player. By means of a local area network the workstations and the personal computers are connected to a VAX 4500 mini computer. All of the data of the *VMNW* is stored in this computer: the corpus texts, the material files created during the editing process and the Lexical Database.

SOFTWARE

The editors work with a fully computerised system, called *Malegys*. During the years 1988-1991, this system was built according to the editors' wishes at the Instituut voor Nederlandse Lexicologie by the E.D.P. department led by P.H.J. van der Kamp. The name Malegys stands for *material* (MA), *lexicography* (LE) and *Gysseling* (GYS), the latter being the editor of the *Corpus*, Dr M. Gysseling. *Malegys* also happens to be a surname in our material: Item van janne

malegyse v lb die hi leende der stad, *Corp.I* 1007,34 (Dordrecht 1284-1285).

The system consists of three parts: 1) the corpus database, containing the electronic version of the source material; 2) the lexicographer's workbench, where the editors analyse and interpret the data derived from the corpus database, and 3) the lexical database which is, in fact, the electronic version of the *Vroegmiddelnederlands Woordenboek* in statu nascendi. For more details about these three components, see the contribution of T.H. Schoonheim, elsewhere in this book.

The operating system is open VMS 6.0 and Oracle RdB is used as a relational database management system. The PCs run MS DOS 6.22 and MS Windows 3.11. They are also equipped with eXcursion, an Xserver by DEC. This software enables the editors to use the same graphical-user interface on the VAX VS 3100 workstations. However, the *Malegys* system itself has a character-based interface.

Special Remarks

NAMES

The official documents contain a lot of names, both anthroponyms and toponyms. Many of these names date from the pre-Dutch, or West-Germanic, period. However, surnames were also often formed with elements of the contemporary language. The many cultivations and reclamations of land in the twelfth and thirteenth centuries resulted in new toponyms, also formed with elements of the contemporary language. Not incorporating those names in the dictionary would deprive us of very relevant language material. The dictionary also describes first names, because of the many possibilities of variation which these names reflect, often clearly following the Middle Dutch language-system, like forming a diminutive by adding *-in, -lin, -kin.* (The same applies for surnames which take suffixes for collectivisation (*-inc*) or feminisation (*-inne*)).

LATIN/MIDDLE DUTCH

In thirteenth-century writing, Latin was gradually replaced by the vernacular. The oldest documents in the Corpus, being Latin texts

containing some Dutch words, are a reflection of this gradual transition.

A Middle Dutch counterpart was not immediately available for every word. This is especially true for names of (exotic) plants and animals. As a matter of fact, up until now many of those words are still not replaced by an indigenous word. Viewed in that light, the editors think it is justifiable to incorporate those originally Medieval Latin words in the *Vroegmiddelnederlands Woordenboek*, e.g. *hiena*, *rinoceros* or *margarita* ('pearl').

On the other hand, there were scribes who, for some technical terminology, appear not to have known the exact Latin counterpart; the word probably did not even exist in Latin. Therefore some Dutch words were latinised, received a Latin-like appearance, e.g. *werpire* < *werpen* (to throw). The same happened with toponyms, e.g. *hollandia* < *hollant* (Holland). This does not only hold true for Latin texts containing Dutch words or words from Dutch origin, but also for notations *in dorso* that are sometimes a mixture of both Latin and Dutch or latinised Dutch.

SAMPLE ENTRY

somer (I) znw.m.,v. — headword, word class
zomer; Zomer [90] — definition, number of occurences

(somer 1e art.) 1/54/35 Gent 1236. soemer, somer, sumer, smere, zoemer, zomer; *ns* -; *d/as* -, -e; *gs* -s; *ds* -e, *ook in de vorm* somre *met syncope van vocaal*; *as* -. *Te Voorne (Zeeland) werd* in die sumer *aangetroffen; dit wijst op een vrouwelijk genus.* — survey entry in *MNW*; frequency per genre: vocabularies/charters/literary texts; oldest attestation (place, year); spelling variants; inflection; morphology, etymology and/or regional variation

I 1. Zomer. || estas : somer, *Bern. 193,4 (Limb 1240).* Ende alle als si uasten bede gandse en*de* sieke mogen hebben den orlof dien de heilege kerke es gewone tegeuene anderen kerstinen liden in den somer. *Corp.I 25,21-23 (Gent 1236).* Het was in enen somer clare Jnt alre schoenste van den ijare (...), *Lutg.K 439,28-29 (Wbra 1265-1270).* Het was op en*en* somer schone Dat liliblomen sijn gewoene Te wassene. (...), *Lutg.K 492,14-16 (Wbra 1265-1270). (...)* — Division (I), definition (1), || beginning of quotation block; quotations: text, source text (page, line), place and year of attestation

In de verb. (des) somers. **In de zomer.** || tes some*r*s caent he*m (t.w. de eekhoorn)* so begome*n* dat heuet te winte*r* spise alle boe*m* urucht va*n* soete*r* wise, *Nat.Bl.D 113,9-11 (Wvla 1287).* sijn smare *(t.w. van de vis trebius)* wonde*r* so es dit somers swart *ende* winte*r*s wit, *Nat.Bl.D 269,31-32 (Wvla 1287). (...)* — collocation: text, definition, quotations; (editorial comment)

1.1. In dateringen ter aanduiding van het feit dat de genoemde dag in de zomer valt. *Aangetroffen bij (de dag voorafgaand aan) de feestdagen van de H. Basilius (14 juni), de H. Johannes de Doper (24 juni), de H. Petrus (29 juni), de H. Thomas (3 juli) en de H. Martinus (4 juli).* || saterdaghes vor sint johans daghe jn middel somer. *Corp.I 139,5-6 (Brugge 1269).* tsfrin*en*daghes na sinte ihans daghe middels some*r*s, *Corp.I 288,37 (Brugge 1275).* van seinte jehans dage te haluen smere, *Corp.I 371,31 (Tongeren? 1277). (...)* — additional remarks

CD-ROM Middelnederlands

CD-ROM Middle Dutch

Information about the Project

INSTITUTE

Instituut voor Nederlandse Lexicologie

ADDRESS

Information can be obtained from:
D. Geirnaert
Instituut voor Nederlandse Lexicologie
Postbus/PO Box 9515
NL 2300 RA LEIDEN
The Netherlands

e-mail: Geirnaert@rulxho.leidenuniv.nl
tel: +31 (71) 527-2283
fax: +31 (71) 527-2115

Contents of the CD-ROM

GLOBAL

The *CD-ROM Middelnederlands* will be made up of a lexicographical and a textual component. The former will correspond to an electronic version of the 11-volume *Middelnederlandsch Woordenboek* (*MNW*, the Dictionary of Middle Dutch); the textual component will contain 1) the two series of the *Corpus van Middelnederlandse teksten (tot en met het jaar 1300)*, a collection of strictly diplomatic editions of all thirteenth-century texts in Middle Dutch and 2) a large amount of other Middle Dutch texts, for the major part based on editions, already existing in print.

DETAILED

The *MNW* (period of publication: 1882-1929) is an alphabetically ordered, scientific dictionary, built on the nineteenth-century linguistic and philological principles that also produced works such as the *Oxford English Dictionary* or the *Deutsches Wörterbuch*. The *MNW* describes the Dutch vocabulary of the middle ages, covering thus the period of c. 1200-1500; the area described coincides mainly with the present-day Dutch-speaking area (the Netherlands and Flanders in Belgium).

Every semantic, syntactic or grammatical distinction made is illustrated with citations, taken from a large collection of source texts. These texts are literary as well as non-literary, although the literary are predominant by far. A detailed description of the editions, manuscripts or printed books used as sources can be found in the volume *Bouwstoffen* ('the materials'), an alphabetically ordered and continuously numbered list of some 1500 items; in this way, this volume is indispensable, e.g. to add information on location and dating to the citations in the other volumes.

The electronic version of the *MNW* will be divided into three electronic "books" or files, each of them separately appropriate for consultation: 1) the dictionary as such (ten volumes), 2) the *Bouwstoffen* and 3) the complete collection of citations used in the *MNW*, adapted to be dealt with as an independent data file ("citation book"). Links between these three parts will be supplied.

For further details on the *Corpus van Middelnederlandse teksten (tot en met het jaar 1300)* one is referred to the *Forschungsbericht* on the *Vroegmiddelnederlands Woordenboek (VMNW)*, elsewhere in this book, where all necessary information on these texts is given extensively. As far as the other Middle Dutch texts of the textual component are concerned, the final selection will be completed in spring 1997); nevertheless, it is already safe to say that in any case all important texts and types of the Middle Dutch literary canon will be present, as well as the most interesting *artes* texts and writings in the religious domain.

The textual component too will be built up of several electronic "books": 1) the first series of the *Corpus van Middelnederlandse teksten (tot en met het jaar 1300)* (c. 2000 official documents); 2) the second series of this *Corpus* (the literary and *artes* texts), and 3) the remaining Middle Dutch texts, divided into verse and prose.

It will be possible to use each of the two components on its own: the lexicographical component as a dictionary and a citation bank, the texts as a corpus for several kinds of lexical, morphological, syntactical and literary research. It is, however, also the intention to create a practicable bond between the two components. This can be realised by filling up the textual component in the first place with source texts of the *MNW*, and by providing the possibilities for a user to switch from dictionary to text. This switching will enable him or her to look up the context of a citation, or, by compiling a greater amount of textual data, to nuance, complete or even correct a certain distinction made.

EXAMPLES OF SEARCH OPERATIONS

Some general applications of the electronic *MNW*: searching the headword index, the full text index and the retrograde index; looking for a particular word-combination, author or source; adding one's own remarks to citations or entries. As for the textual component: consulting every single text as well as a self-selected corpus of texts; searching an alphabetical and retrograde index; looking for word-combinations; adding one's own remarks and annotations.

History of the Project and Planning

SHORT HISTORY OF THE PROJECT

In 1995 the *Instituut voor Nederlandse Lexicologie (INL)* issued a CD-ROM of the 37 volume *Woordenboek der Nederlandsche Taal* (*WNT*, the Dictionary of the Dutch Language), the Dutch counterpart of the *Oxford English Dictionary*. This scientific dictionary, the source for all other Dutch dictionaries, gives a detailed description of all Dutch words from the sixteenth to the twentieth century, illustrating the meanings and the use of a word with citations taken from works of that period. This release on CD-ROM meant an important expansion of the uses of the *WNT* and it was consequently obvious that an electronic version of the *Middelnederlandsch Woordenboek*, the dictionary preceding the *WNT*, would be very welcome too. Moreover, this accorded with the *INL* intentions to use amongst others files such as the electronic *WNT*, *MNW* and *VMNW* (the Dictionary of Early Middle Dutch; see elsewhere in this book) to build an integrated language database of the Dutch language from the twelfth till the twenty first century.

NWO, the Dutch Organisation for Scientific Research, was willing to grant *INL* a subsidy for this Middle Dutch project, on the following conditions: 1) a corpus of Middle Dutch texts should be added to the dictionary; 2) the CD-ROM should be available mid-1998 at the latest, and 3) the price for the CD-ROM should be sufficiently low to reach a target group as large as possible.

PROGRESS AND PLANNING

Currently the proceedings of the project are on schedule. The electronic *MNW* (with the exception of the "citation book") and the *Corpus van Middelnederlandse teksten (tot en met het jaar 1300)* are virtually ready and already available on a trial CD-ROM. In the coming months the main activities will be: 1) refining the citation book of the *MNW*; 2) establishing the definite list of the Middle Dutch texts to be included in the textual component; 3) the uniformisation of the texts already present in digitised form and 4) the acquisition and uniformisation of the digitised versions of the texts that are still lacking.

Technical Information

The technical part of the project is done by the Rotterdam firm AND Electronic Publishing B.V., which also realised the *WNT* CD-ROM. The application to access the different electronic books is AND CompLex 3, a newer version of the one used for the *WNT* CD-ROM.

System requirements for the Windows version:

◊ an IBM or fully compatible PC with 80386 DX processor or higher, 20 Mhz or faster (486 recommended);
◊ Microsoft Windows 3.1 or Windows for Workgroups (3.11) with Win32s 1.3 (supplied with the product) Windows 95, Windows NT3.51, or Windows NT4.0 (or higher);
◊ 8 Mb of RAM recommended;
◊ 4 Mb free hard-disk space (or 15 Mb if Win32s needs to be installed);
◊ 256-colour (S)VGA colour monitor;
◊ an MPC-rated 150 Kb/s CD-ROM drive (or higher transfer-rate);
◊ a Windows-supported mouse or pointing device;
◊ a Windows-supported printer with associate drivers;
◊ Ancillary Software requirements: MSCDEX 2.21 or higher (not supplied with the product).

Figure 1. The search for the word artur, selected in the full text index of the *MNW*, produces 133 hits (= 133 headwords where this form can be found at least once); the illustration shows a) the result in the headword achtergaen; b) the result window and c) the "toolbar", a bar in the main window with buttons representing the main functions of AND CompLex 3.

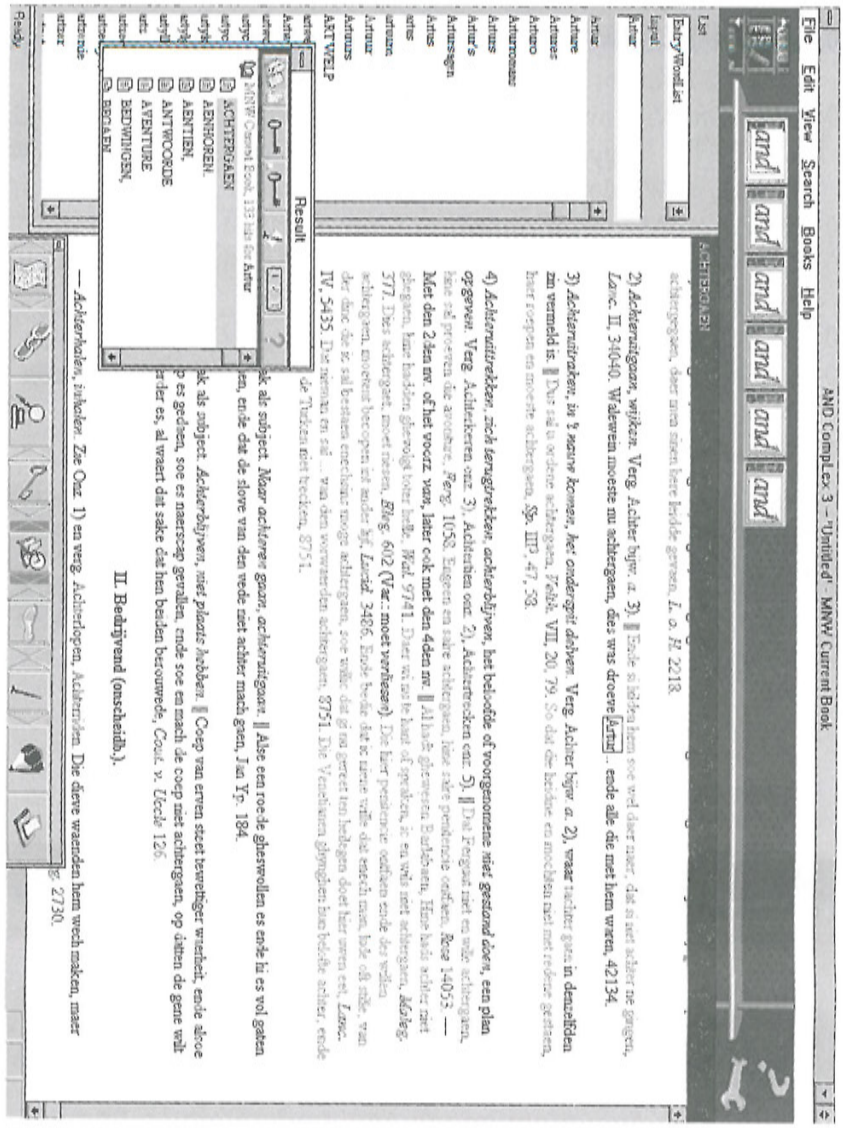
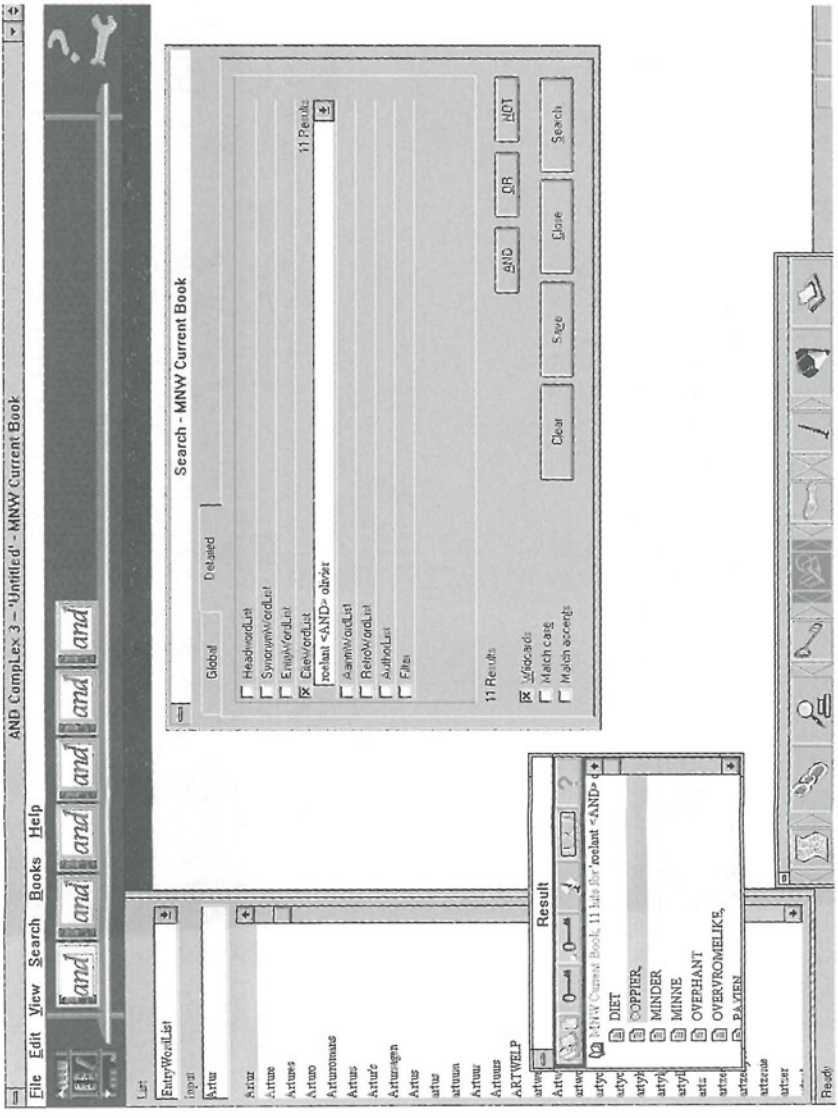


Figure 2a/2b. With the "detailed search dialog" and the use of Boolean operators, it is possible to look for word combinations in the different indices; a search for the combination roelant and olivier in the file with only the citations of the *MNW* gives 12 results; the illustration (2b) shows the result in the headword coppier.

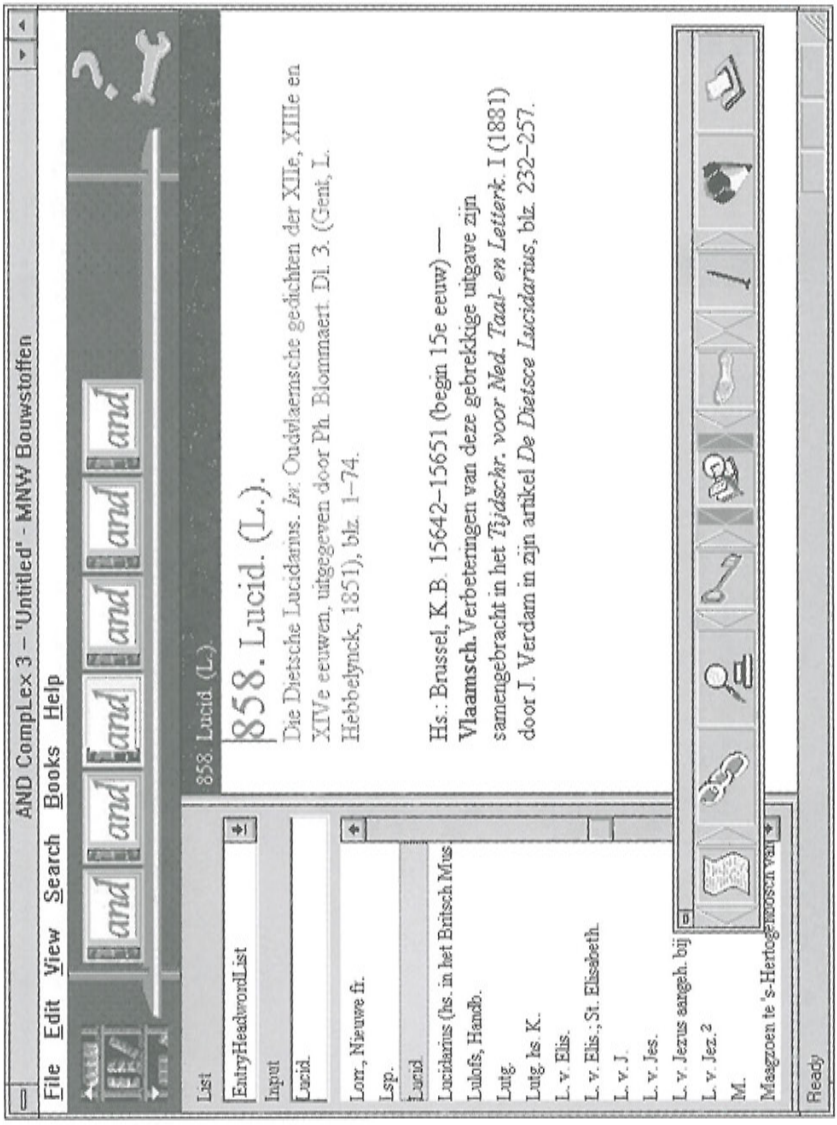
Figure 2a



AND CompLex 3 – 'Untitled' - MNW Current Book
File Edit View Search Books Help
and and and and and and
COPPIER,
COPPIER, znw. m. Misschien het mnl. coppe in de bet. *haan*, met een rom. uitgang. Vgl. *Vrouw. e. M.* Gloss. bl. 139. || (Hi) recht den hals als een coppier: Roelant ... noch Olivier en hadden niet bi hem een spier, *Vrouw. e. M.* VII, 229.
Ready

Figure 2b

Figure 3. One of the "electronic books" on the CD-ROM contains the Bouwstoffen, the list of sources for the *MNW*; entering or clicking one of the numbers (or titles) of the list, one gets the article as it is found in the book; a full text index can be used too.


AND CompLex 3 – 'Untitled' - MNW Bouwstoffen
File Edit View Search Books Help
and and and and and and
List
EntryHeadwordList
Input
Lucid.
Lorr., Nieuwe fr.
Lsp.
Lucid
Lucidarius (hs. in het Britsch Mus.
Lulofs, Handb.
Lutg.
Lutg. hs. K.
L. v. Elis.
L. v. Elis.; St. Elisabeth.
L. v. J.
L. v. Jes.
L. v. Jezus aangeh. bij
L. v. Jez.2
M.
Maagzoen te 's-Hertogenbosch van
858. Lucid. (L.).
858. Lucid. (L.).
Die Dietsche Lucidarius. In: Oudvlaemsche gedichten der XIIe, XIIIe en XIVe eeuwen, uitgegeven door Ph. Blommaert. Dl. 3. (Gent, L. Hebbelynck, 1851), blz. 1–74.
Hs.: Brussel, K.B. 15642–15651 (begin 15e eeuw) — Vlaamsch. Verbeteringen van deze gebrekkige uitgave zijn samengebracht in het Tijdschr. voor Ned. Taal- en Letterk. I (1881) door J. Verdam in zijn artikel De Dietsce Lucidarius, blz. 232–257.
Ready

AND CompLex 3 – 'Untitled' - Corpus Gysseling I Ambtelijke Bescheiden
File Edit View Search Books Help
and and and and and and
List
Artikelen
Input
0213
0213 0214 0215 0216 0217 0218 0219 0220 0221 0222 0223 0224 0225 0226 0227 0228 0229 0230 0231 0232 0233 0234 0234a
0213
383
213 Gent 1277 december 19
Schepenen van Gent oorkonden dat Symoen van den Pitte, Jan de Jonghe, Juris Brebart en Jacob van den Wallekine, poorters te Gent, erkennen samen 62 pond sterling schuldig te zijn aan Jan van Wickede of aan brenger dezes, te betalen op Iepermarkt eerstkomende.
A. Origineel: Gent, Rijksarchief, fonds Gaillard, nr. 756. Perkament (H 68, B 207 mm). Chirograaf; onderaan het devies MAGN.
Het si cont al den ghenen die dese lettren sullen horen ende zien dat der symoen van den pitte ver lucien sone. ende jan de jonghe. ende juris brebart. ende jacob van den wallekine porters in ghent sin sculdech. helc alse barblikenst van hem vieren den her janne van wickede .lxii. lib esterling. ofte soe wien soe desen middelen brief bringt. te gheldene binnen jpermarct dat naest Orconde den her janne ser jans sone ser sanders sone. ende den her boidine ser pieters sone. alse schepenen. Dit was ghedaen jnt jaer van den jncarnatione ons heren .M°. cc°. lxx° vii°. Sondaghes vor kersauonde. Jn te meester danels.
Ready

Figure 4. The text of one of the charters in the *Corpus van Middelnederlandse teksten (tot en met het jaar 1300)*.

Figure 5. Using the retrograde index of the *Corpus van Middelnederlandse teksten (the literary manuscripts)*, one can easily collect, e.g., all words ending in a particular character or suffix. The illustration shows the index representing all words ending in -atie and the textual place of the word ymaginatie.

Figure 6a/6b. Here a combined search is made in four indices of the *Corpus van Middelnederlandse teksten (the official documents)*: a) shows the window for the search for a person named jan mont in the Middle Dutch text of all documents from Bruges, written in the period 1275 - 1285; b) gives the first result.

Figure 6a

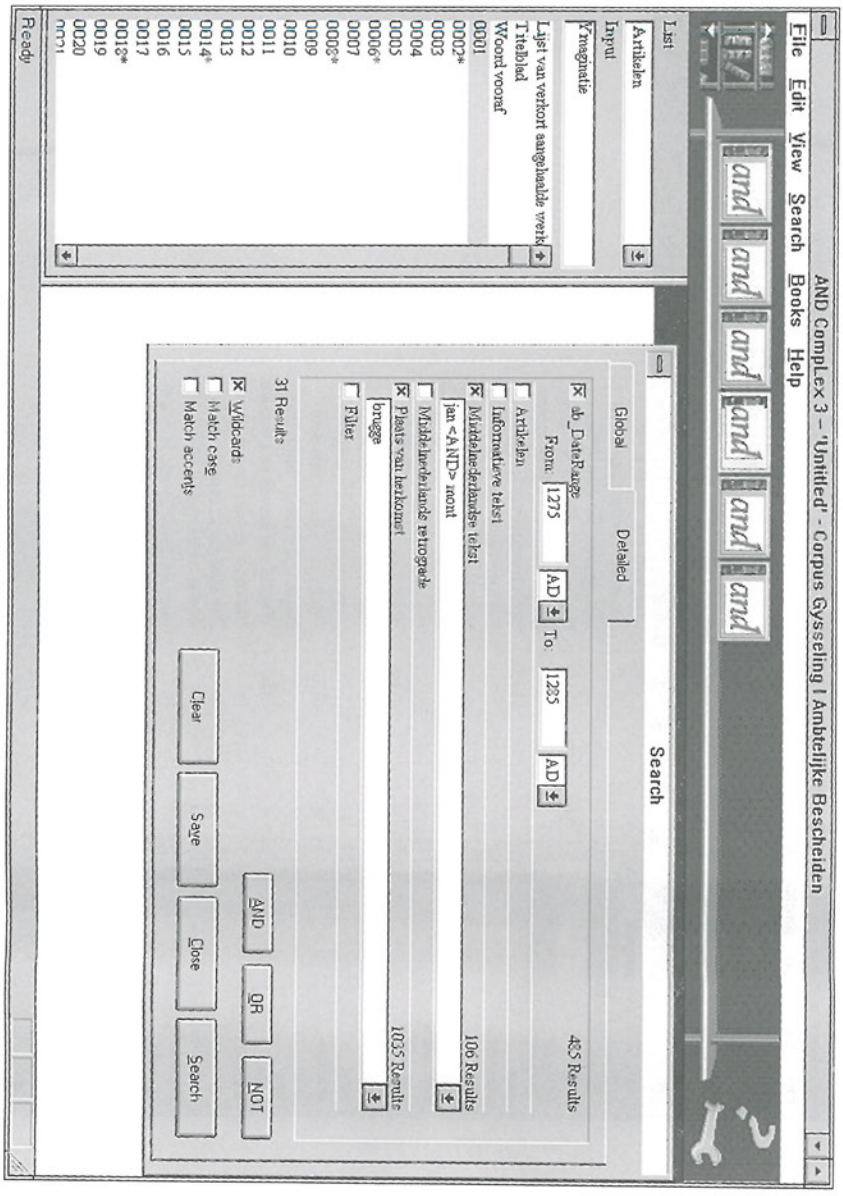


Figure 6b

AND CompLex 3 – 'Untitled' - Corpus Gysseling I Ambtelijke Bescheiden

File Edit View Search Books Help

and and and and and and

0249

249 Brugge 1279 mei 3

Schepenen van het Vrije oorkonden dat Alard Hereganc 4 lijnen land te Schoondijke voor 9 jaar in pacht nam van Sint-Salvators te Brugge, tegen 22 s. 9 d. 2 kapoenen 's jaars.

A. Origineel: Brugge, Bisschoppelijk Archief, Sint-Salvators. Perkament (H 163 + plica 9, B 160 mm); de 2 uithangende zegels verloren. In dorso: (4e kwart 13e e.:) alard hereganc xxii so&bstrok; ix đ ii capones.- (idem:) jn sconendijc.- (15e e.:) vaga.

Deze oorkonde is van dezelfde hand als nr. 240.

Clais f domini hughes jan vanden zande. hughe van boyenghem. jan mont & bouden van laepscure wie scepenne vanden vrien doen te verstane allen den goenen die dese lettren zullen zien jof horen lesen. dat cam vor ons alard hereganc & waerd wettelic censere dese naeste .neghen. jare die nu comen. sere gherards van bredene capellaen van sinte saluators in brugghe ter vorseider kerke boef van. viere linen lands lichtelic min jof me legghende binden ambochte van oestburgh indie prochie vanden sconendyc of nordalf hughe pra segers lant of sutalf gillis f symons west over den wech die welke die gaet te lambays waerd al dit vorseide land omme .xxij. &stilde; & .ix. d goeder vlamscher

Ready

Historisch Woordenboek der Nederlandse Rechtstaal 14de-18de eeuw

Historical Dictionary of
Dutch Legal Terminology
14th-18th Century

Information about the Dictionary

EDITORS

Project Leaders:	Em. Prof. Dr R.C. Van Caenegem
	Prof. Dr J. Monballyu
Research Assistant:	C.M.O. Verhas, M.A.

INSTITUTE

Project funded by the *Nationaal Fonds voor Wetenschappelijk Onderzoek* (National Fund for Academic Research), affiliated to the *Koninlijke Academie voor Wetenschappen, Letteren en Schone Kunsten van België – Klasse der Letteren* (Royal Academy for the Sciences, Arts and Fine Art of Belgium – Class of the Arts).

ADDRESS

HWNR-Project
C.M.O. Verhas
Paleis der Academiën
Hertogsstraat 1
B 1000 BRUSSELS
Belgium
e-mail: Christel.Verhas@rug.ac.be
tel.: +32 (2) 550-2349
fax: +32 (2) 550-2335

Contents of the Dictionary

TYPE

The *Historisch Woordenboek der Nederlandse Rechtstaal (HWNR)* is an alphabetically ordered, scientific, diachronic and explanatory dictionary. Each entry listed is accompanied by information on grammatical status, variance of spelling, etymology and different meanings. The latter is illustrated with examples from legal texts, in which the relevant term is given in the context used. Nouns and pronouns are listed in the nominative singular, verbs in the infinitive.

SUBJECT/CORPUS

The dictionary is based on a corpus of legal texts written between 1300 and 1800, supplemented with information taken from glossaries of old legal terms. The sources used are directly linked with the four official cornerstones of the law: legislation, common law, justice administration and jurisprudence. The body of texts upon which the dictionary is built is not uniform in nature: it encompasses hand-written and printed ordinances, rules of common law and trial transcripts. Given the full extent of potential sources, it is necessary to be selective, whilst ensuring that the history of Dutch legal terminology is still reflected accurately from both geographical and chronological points of view.

DESCRIBED PERIOD

The "Corpus-Gysseling" incorporates all Middle Dutch texts predating the fourteenth century and is therefore the main source for the *Vroegmiddelnederlands Woordenboek (VMNW)* (Dictionary of Early Middle Dutch). This also includes Early Middle Dutch legal terminology. The *Historisch Woordenboek der Nederlandse Rechtstaal* follows on chronologically from the *VMNW*. The earliest sources used date from the fourteenth century.

With the introduction of the *Code Napoleon* at the end of the eighteenth century, the legal practices and terminology which had been used in the Netherlands prior to this date were fundamentally reformed. With the exception of those glossaries and jurisprudential

texts which are of historical relevance, the most recent sources date from the eighteenth century.

DESCRIBED AREA

Selection criteria must also be enforced when it comes to geographical considerations. In some regions of what is now the Netherlands and Flanders, languages other than Dutch were adhered to. The texts upon which the *HWNR* is based are derived from the Dutch-speaking regions of the old Netherlands: Flanders, Brabant, Holland, Zeeland, Utrecht, parts of Guelderland and Limburg. The Frisian and Ancient Low German which was spoken in some parts of the Netherlands will not be treated here in the *HWNR*. However, the Dutch legal terminology used in these areas will, where possible, be incorporated.

EXAMPLE OF AN ENTRY

After every entry there follows a short grammatical interpretation and summary of possible deviations in spelling. Then follows a review of the various meanings attached to the entry, according to those found in the body of texts. The contexts in which the entry was used will be reported chronologically, with references to the sources from which the quotations were extracted.

History of the Project and Planning

SHORT HISTORY OF THE PROJECT

Professor F.L. Ganshof had already proposed the preparation and publication of an historical lexicon of Dutch legal terminology as early as 1941, in the Class of the Arts of the Royal Flemish Academy. However, it was not until 1992 that his proposal and its desirability were discussed again, this time by the Academic Committee for the History of Law. Following these discussions, its president, Professor R.C. Van Caenegem contacted The Royal Dutch Academy for the Sciences and the Institute for Dutch Lexicology. In 1993 he presented his report to the Academic Committee for the History of Law. A proposal for the compilation of a concise dictionary, in which existent material would be amassed

and supplemented with new source-based research, would be presented to the Academy. In 1994, Em. Professor G. Verbeke, Permanent Secretary of the Royal Academy for the Sciences, Arts and Fine Art, announced that the *HWNR*-project would commence as of 1 January 1995. Prof. W.P. Gerritsen, Chairman of the Dutch Academy's Department of Literature, was requested to co-operate. 1 June 1995 marked the definite start of the research project, "Edition of a comprehensive historical dictionary of Dutch legal terminology from the thirteenth century to c. 1800". The research is funded by the National Fund for Scientific Research whilst the Royal Academy in Brussels will receive the research findings. Prof. Dr J. Monballyu was appointed as project leader; Professors Van Caenegem and Verbeke are associated with the project in their capacity as supervisors.

The initial phase was realised by M. Oosterbosch M.A. He conducted research into the feasibility and potential of the original project goals and presented a rough outline of the research methodology. Given that co-operation with the Netherlands is yet to be fully realised, the collection and compilation of the sources in this preliminary phase is restricted to legal sources from the southern parts of the old Netherlands. After the departure of M. Oosterbosch M.A., the research was continued by C. Verhas M.A.

YEAR OF PUBLICATION

Given the recent launch and extent of the research project, it is too early at this stage to supply a date of publication.

FORM OF PUBLICATION

The *Historisch Woordenboek der Nederlandse Rechtstaal* will probably be published in print and on CD-ROM. Ideally, the body of texts on which the dictionary is based will also be available on CD-ROM.

Lexicographical Tools and Methods

HARDWARE

At the moment the project has at its disposal a Digital Venturis Pentium-60 M540 with a 15" screen and CD-ROM, an HP Laserjet 5P printer with parallel cable and an HP Scanjet 3P with software applications. The Dutch Historical Data Archives in Leiden is kindly offering its assistance with the scanning of old source-publications onto CD-ROM.

SOFTWARE

In addition to the Software (Novell WP 6.1 and Microsoft Access 2.0 for Windows) which is currently being used for the correction of digitised texts and word-processing, the project will, in co-operation with the Institute for Dutch Lexicology, seek to obtain a programme which will yield the best results in the processing of entries.

Special Remarks

With regards to old-fatherland law, one cannot deny the influence of Romano-canonical law and old French judicial practices. This has partly resulted in the fact that Dutch legal terminology is tinged with words of Latin or French origin, but which are used in the texts in quasi-Dutch form (e.g. *mainteneren, interloqueren,...*). Because these words are clearly part of Dutch legal terminology in the Ancien Régime, they are also listed in the dictionary.

A Thesaurus of Old English

Information about the Thesaurus

EDITORS
Jane Roberts and Christian Kay
with Lynne Grundy

INSTITUTES
University of Glasgow & King's College London

ADDRESS
Correspondence to be sent to:
Jane Roberts
English Department
King's College
Strand
London WC2R 2LS
UK

e-mail: J.Roberts@kcl.ac.uk
tel.: +44 (171) 873-2181

Contents of the Thesaurus

TYPE

A Thesaurus of Old English (*TOE*) is a conceptually-ordered period thesaurus. The c. 48,000 entries present the English vocabulary of the Anglo-Saxon period in two ways: word-senses are given in 18 ordered categories (volume I); and an alphabetical index identifies the category or categories in which the word-senses are to be found (volume II). Indications of meaning are given at each level, and the defining headings are written so as to reflect grammatically what is defined.

SUBJECT/CORPUS

The word-senses arrayed in the thesaurus reflect the information to be found in the standard dictionaries of Old English: J. Bosworth and T.N. Toller, *An Anglo-Saxon Dictionary* (London, 1898); T.N. Toller, *An Anglo-Saxon Dictionary: Supplement* (London, 1921); A. Campbell, *An Anglo-Saxon Dictionary: Enlarged Addenda and Corrigenda* (Oxford, 1972); J.R. Clark Hall and H.D. Meritt, *A Concise Anglo-Saxon Dictionary*, 4th edn (Cambridge, 1960). Although during the final stages of preparation of the thesaurus it was possible to use the materials published by the Toronto *Dictionary of Old English* (described elsewhere in this book) project for checking the data, essentially the meanings depend upon the definitions to be found in the older dictionaries. Proper names are generally excluded, and not all words from the language's closed systems are included.

DESCRIBED PERIOD

The thesaurus contains the extant vocabulary of over 500 years of the history of the English language. Very little evidence is extant from the eighth century, rather more from the ninth, much more from the tenth and most from the eleventh and twelfth centuries.

DESCRIBED AREA

The extent of Anglo-Saxon England underwent many alterations. Eighth-century materials can come from as far north as what is today south-west Scotland, where the inscribed Ruthwell Cross is still

to be found in Dumfriesshire, or from the northern England monastic complex at Monkwearmouth and Jarrow, the likely home of one of the earliest copies of Cædmon's Hymn. The most important ninth-century materials in English are associated with King Alfred of Wessex and were written in Wessex and Mercia: the Viking settlement of the Danelaw (the area north of a line running from Chester to London) makes it unlikely that much English is recorded outside the southern part of England for two or three generations from the middle of the ninth century. The northern English of the glossator Aldred, who is likely to have been based at Chester-le-Street in the latter half of the tenth century, shows heavy Scandinavian influence. Overall, however, it seems either that written Old English was conservative or that it comes mainly from the southern part of the country, where the language changed less rapidly.

English was not yet the language of Ireland, Scotland or Wales.

EXAMPLE OF AN ENTRY

Volume I presents the hierarchical classification, which proceeds from the most general terms to the most specific. Thus, each meaning entered is defined not only by its own heading but by the headings above and below it in the structure. Within each heading further subordination or co-ordination may be indicated, as in the following selection:

02.06 Animal
02.06.03 Wild animal
02.06.03.01 Particular animals (alphabetical order)
02.06.03.01.13 Rodents
.Mouse: mūs
.. A thieving mouse: mūsþēofᵒᵍ
.Rat:

If the word we are interested in is *mus*, we find out about it by reading back through the information provided by means of the numerical hierarchy: it is a rodent, that is a particular sort of animal, it is wild, then animal. The unnumbered dots give further information: both 'Mouse' and 'Rat' are subordinate to 'Rodent'

and co-ordinate with each other, while *musþeof*, with two dots, is subordinated by a further degree of specificity. Volume II is an index, listing all the Old English words and phrases in the thesaurus, together with the lowest numbered headings under which they appear, for example:

> **mūs** 02.04.05.05 Muscle; 02.06.03.01.13 Rodents
> **mūsþēof** 02.06.03.01.13 Rodents

This index allows the user to see the range of representation of polysemous words within the classification.

Four symbols are used as annotations of the entries in the thesaurus classification. The flags are superscript, and are placed immediately after words, to indicate frequency. They relate to word forms, not to meanings. Phrases are not marked. Noncewords are flagged **o**, a label that can co-occur with **p** (words restricted to the Old English poetic corpus) or with **g** (words found only in glossed texts or glossaries). The occasional use of the flag **q** points to highly dubious forms. The absence of any flagging therefore indicates that the word appears more than once in Old English and that it is not restricted either to poetry or to glosses.

History of the Project and Planning

SHORT HISTORY OF THE PROJECT

The *TOE* is a pilot study for *The Historical Thesaurus of English* (*HT*). Its first stage was the creation of an archive of slips by Jane Roberts, undertaken to supply the *HT* editors with some idea of the range of Old English vocabulary not to be found in the *Oxford English Dictionary*. For each meaning found in the standard Anglo-Saxon dictionaries a slip was made and assigned an initial sorting number or numbers taken from the 1962 edition of Roget's *Thesaurus*. Once the compilation of this experimental archive of slips was complete, a *Skeleton Old English Thesaurus* was created (Research tool on nine microfiches, prepared by Jane Roberts with Christine Brown at the King's College London Computing Centre,

1982), to establish a record of the word senses excerpted from the standard dictionaries. The *Skeleton Old English Thesaurus* provides both a preliminary Old English thesaurus arranged according to Roget-derived sorting numbers and an alphabetic index to the experimental archive.

Because the *OED* generally omits those words that had dropped out of use by 1150, the skeleton thesaurus provided the *HT* editors with a complementary overview of the earliest vocabulary of English. With the decision to incorporate the ancillary Old English materials into the *OED*-derived *HT*, work towards the *TOE* got under way. The relative smallness of the separate Old English archive (c. 48,000 word senses as opposed to some 650,000 word senses drawn from the *OED* for the *HT*) made it a useful corpus for a full-scale pilot study. Christian Kay and Jane Roberts started from the 26 major categories by then established for the *HT*. That the *TOE* has only 18 categories illustrates the general principle on which both classifications are based: that the structure must derive from the data rather than be imposed upon them. Some of the separate categories of the *HT* set up divisions that seemed impracticable for the Anglo-Saxon period. Thus, for example, the *TOE* does not make a major division between "The Supernatural" and "Institutional Religion", but presents a unified category "Religion". Lynne Grundy joined the editors in the final editing period, with part-time funding from the King's College Research Strategy Fund. Jane Roberts and Christian Kay had a small grant from the British Academy (1984), to assist the beginning of data input in Glasgow. Jane Roberts held a Leverhulme award (1985-86), to enable her to complete the draft arrangement of the separate Old English archive of slips within the framework of the *HT* classification.

YEAR OF PUBLICATION

1995

FORM OF PUBLICATION

A Thesaurus of Old English in two volumes. Volume I: Introduction and Thesaurus. Volume II: Index. King's College London: Centre for Late Antique and Medieval Studies 1995

Lexicographical Tools and Methods

THE HARDWARE

At first, information was fed into a database using dBase II software on Apricot microcomputers and floppy disks containing editorial materials, and corrections were regularly exchanged between Glasgow and King's. In 1989 all the current dBase files were converted at King's and loaded into a new system, using INGRES relational database software. By 1992 the volume of material in electronic form made it necessary to transfer the database from a personal computer to a VAX mainframe computer, which had the benefit of allowing simultaneous access by several people. More recently the database was moved to a UNIX machine set as a "database server", but still using INGRES software. Datafiles are now regularly transferred electronically between King's College and Glasgow over "JANET" (Joint Academic NETwork), the communication system which links all UK research and higher education institutions.

THE SOFTWARE

The *TOE* database is held in a single table in INGRES, a relational database. Each record, whether Old English lexical item or Modern English heading, constitutes a row in the database. There are at present 72,612 rows. Each row has a unique "address", composed of a group reference, sub-category reference and individual number. Other data stored there are to do with part of speech, alphabetical-sort sequence, cross-referencing and simple distributional information. Some further fields record editorial materials, for example the older Roget categorisation, any identified modern or earlier English etymological tie-up, and a comment space where notes are made (e.g. locations for rare words, meanings or usage).

Both the Thesaurus and the Index were initially extracted from the database by a custom-made report designed by Harold Short (Centre for Computing in the Humanities, King's College London). Text is recorded in the database in ASCII, so the Anglo-Saxon characters thorn (þ) and ash (æ) are entered as symbols, and the overline markers of vowel length are indicated by underscores after the vowel. Harold Short's report selects the fields needed, in the

right order, and with the right font characteristics. In order to prepare these materials for publication, a further process was created to make them ready for input to LaTeX typesetting software. The LATEX process produced the camera-ready copy sent to the printer.

Special Remarks

The *TOE* database continues to be a living repository of new research as we continue to work on the vocabulary of Old English, and it will play an important part in the new work going ahead at King's College London towards a Middle English thesaurus.

.A bison's horn: wesendhorn[o]

02.06.03.01.05 Boar: bār, eofor, eoforswīn[o], wildeswīn[o]

See 02.06.02.01.07 Pig

02.06.03.01.06 Deer, hart: heorot
.Of/concerning deer: heorten[o]
.A deerskin: heorþa[g]
.A young deer: hindcealf[g]
..Skin of a young deer: næsc
.A male deer: bucca, bucheort[og]
..A stag: hēahdēor, stagga
..A roebuck: rā, rāhdēor
.A female deer: dā, hind
.An elk: eolh[g]
.A reindeer: dēor, hrān, wilddēor
..A reindeer as decoy: stælhrān[o]

02.06.03.01.07 Fox: fox
.A female fox, vixen: fyxe[o]
..A vixen skin: fyxenhȳd[o]

02.06.03.01.08 Goat: stānbucca[og], wudugāt
.A male goat: wudubucca
.A female goat: rǣge

See 02.06.02.01.05 Domesticated goat

02.06.03.01.09 Hedgehog: hærenfagol/hattefagol[q], īl
.Hedgehogs: erinaces[op]

02.06.03.01.10 Marten: mearþ

02.06.03.01.11 Mole: wand[g], wandeweorpe

02.06.03.01.12 Otter: otor

02.06.03.01.13 Rodents:
.Beaver: befer
.Dormouse: hearma[g], sisemūs[g]
.Hare: hara
.Mouse: mūs
..A thieving mouse: mūsþēof[og]

02.06.03.01.14 Weasel

.Rat: ræt[og]
.Shrew: scirfemūs[og], scrēawa[g]
.Squirrel: ācweor(na)[g]

02.06.03.01.14 Weasel: wesle

02.06.03.01.15 Wolf: mearcweard[op], wulf
.Head of a wolf: wulfes hēafod
.A female wolf: wylf[g], wylfen
.Of/concerning a female wolf: wylfen

02.06.04 Exotic animals:

02.06.04.01 Particular exotic animals (alphabetical order):

02.06.04.01.01 Ape: apa

02.06.04.01.02 Camel: camel, olfenda/-e
.A camel as mount: stēda
.A female camel: olfendmyre[o]

02.06.04.01.03 Elephant: elpend, ylp
.An elephant's tusk: elpendbān, elpendtōþ[og]
..Of ivory: elpendbænen

02.06.04.01.04 Hippopotamus: nicor, sæhengest[g]

02.06.04.01.05 Hyena: nihtgenge[g]

02.06.04.01.06 Ibex: firgenbucca[o], firgengāt

02.06.04.01.07 Ichneumon: næderbita[g]

02.06.04.01.08 Leopard: lafor[o], lēopard[o], pardus

02.06.04.01.09 Lion: lēo, lēona/-e
.Flesh of lion: lēonflæsc[o]
.A young lion: lēonhwelp[og]

02.06.04.01.10 Lynx: lox

02.06.04.01.11 Panther: palþer[o], pandher[op]

(ge)mundian 11.10.01 Protection, safekeeping
mundiend 11.10.01 Protection, safekeeping
mundlēow 04.06.01.03 A bath
mundrōf 05.08 Strength
mundwist 11.10.02.02.01 Care, interest in
gemuning 06.01.04 Faculty of memory
munt 01.01.02.01.02.02 Hill, mountain; 01.01.02.01.02.02.02 Mountain
muntælfen 16.01.03.04 Elfin race
muntclȳse 01.01.02.01.03 Hollow/depression in land
muntgīu 01.01.02.01.02.02.02 Mountain
muntland 01.01.02.01.02 High land
munuc 16.02.03.03 A religious; 16.02.03.03.02 A monk; 16.02.03.03.03 A nun
munucbehāt 16.02.03.03 A religious
munuccild 02.03.01.04 Child;16.02.04.07.05.01.01 Going into a monastery, admission
munuccnapa 16.02.03.03.02 A monk
munucgegyrela 16.02.05.07 Monastic garb
munuchād 16.02.03.03 A religious
munuchēap 16.02.03.03.02 A monk
munucian 16.02.04.07.05.01.01 Going into a monastery, admission
munuclic 16.02.03.03 A religious
munuclīce 16.02.03.03 A religious
munuclīf 16.02.03.03 A religious; 16.02.05.04 A monastery/convent
munuclīf dōn/lǣdan 16.02.03.03 A religious
munucrēaf 16.02.05.07 Monastic garb
munucregol 16.02.03.03 A religious
munucscrūd 16.02.05.07 Monastic garb
munucstōw 16.02.05.04 A monastery/convent
munucþēaw 16.02.03.03 A religious
mūr 04.05.02.02 A wall
murcen 08.01.03.04 Grief
murcian 08.01.03.04.01 Complaint, lamentation
murciende 08.01.03.05 Murmuring, complaint
murcnere 08.01.03.04 Grief
murcnian 08.01.03.05 Murmuring, complaint
murcung 08.01.03.04 Grief; 08.01.03.05 Murmuring, complaint
murnan 06.01.08.06 Fear
murnan æfter 06.02.05.01 Strong liking for, devotion to
murnan (æfter, for) 08.01.03.04.01 Complaint, lamentation
murnende 08.01.03.04 Grief
murnung 08.01.03.04 Grief
mūs 02.04.05.05 Muscle; 02.06.03.01.13 Rodents
muscelle 04.01.02.01.02.07 Fish
muscel(le) 02.06.06.04.01.04 Mussel
mūsepise 02.07.11 Plants/flowers (alphabetical order)
mūsere 02.06.08.04 Bird of prey
mūsfealle 04.06.02.02 Foul, filthy, squalid
mūsfealu 03.01.14.09 Dull brown
mūshafoc 02.06.08.04 Bird of prey
must 04.01.03.05.01.01 Wine
mustflēote 02.06.09.02.07 Fly
mustwyrm 02.06.09 Insect/small creature
mūsþēof 02.06.03.01.13 Rodents
mūtung 15.04.01 Lending of money
mūþ 02.04.03.01.03 Face; 02.04.03.01.03.06 Mouth;

A Thesaurus of Middle English

Information about the Thesaurus

EDITORS
Jane Roberts and Louise Sylvester

INSTITUTES
King's College London and University of Glasgow

ADDRESS
Correspondence to be sent to:
Louise Sylvester
English Department
King's College
Strand
London WC2R 2LS
UK

e-mail: Louise.Sylvester@kcl.ac.uk
tel.: +44 (171) 873-2558

Contents of the Thesaurus

TYPE

The proposed *Thesaurus of Middle English* (*TME*) is a conceptually-ordered thesaurus for English from 1100 to 1500. The word-senses in use in English across these centuries are to be presented in some 20 ordered categories, together with the dates recorded for usage. At each level in the classification, indications of meaning will be presented, and the defining headings will be framed in such a way as to reflect grammatically what is defined.

SUBJECT/CORPUS

The nucleus for the collection will be drawn from the files of the *Historical Thesaurus of English* project (*HT*) and therefore reflects the information to be found in the *Oxford English Dictionary* (*OED*). We hope to supplement this corpus from the *Middle English Dictionary* (*MED*), from other specialised work on Middle English vocabulary and from the *TOE*.

DESCRIBED PERIOD

The period covered by the thesaurus begins notionally at 1100, i.e. a half-century before the date from which *OED* materials are comprehensive. (Generally Old English vocabulary is represented in the *OED* only if in continued use after 1150). Thus, supplementation of the *OED* nucleus may be particularly marked at the beginning of the designated period.

DESCRIBED AREA

The use of English spread almost throughout Britain and into parts of Ireland in the Middle English period. The area described is therefore progressively enlarging, with Welsh, Irish, Scots Gaelic and the Scandinavian dialects of the Western Isle in retreat.

EXAMPLE OF AN ENTRY

There is as yet no agreed sample entry. The sample supplied, taken from Louise Sylvester, *Studies in the Lexical Field of Expectation* (Rodopi, 1994), incorporates at this stage no new information.

History of the Project and Planning

SHORT HISTORY OF THE PROJECT

The *TME* follows on from the *TOE* project, for which the final camera-copy was prepared at King's College London. With both the *HT* and *MED* nearing completion, it seemed time to undertake sustained research in preparation for the *TME*. Preliminary discussion began in 1994, and work on the related lexical studies volumes got under way late in 1995, when Louise Sylvester took up a British Academy Institutional Fellowship based in the English Department at King's College.

The preliminary project has the short-term aim of producing two volumes of Middle English lexical studies. The first of these, "Middle English Word Studies", is to contain detailed bibliography specifically about Middle English lexicology. The second volume, "Middle English Semantic Field Studies", will describe and analyse studies of Middle English lexical fields. The first of these volumes, to be completed in 2000, will make available a register of where significant work on word meaning is available, and the files from which it will be drawn contain the details of the relevant discussion necessary both for the second volume and for the *TME*. Work proper on the *TME* will begin when all the archive materials at Glasgow have been edited into the *HT* database.

PLANNED YEAR OF PUBLICATION

The *HT* archive materials at Glasgow will all be in the machine-held database in 2000, as the first of the King's College lexical studies volumes is completed. At that point the database materials for the word-studies volumes should be a valuable tool in the supplementation of the *HT* (*OED*) dates . Reading the downloaded materials against the *MED* will set a new challenge. The process will run alongside and complement the completion of the volume on Middle English lexical field studies. The *TME* is expected to be available in 2005.

FORM OF PUBLICATION

The *TME* will be published in print.

Lexicographical Tools and Methods

THE HARDWARE

At this point we envisage the use of hardware and software analogous to that used both for the *HT* and the *TOE*.

THE SOFTWARE

For the Middle English lexical studies project we currently use a PAPYRUS database.

EX COURAGE

X.1 COURAGE

The quality of courage

(nouns) *ellen*/elne OE-a1240; courage 1375+1382—; coraiouste 1382 *obs, rare*; inwit 1382; courageousness 1494-1605

Courageous

(adjectives) courageous 1297—; hearty c1380-1684; leonine c1386-1822

In a courageous manner

(adverbs) courageously 1477-1847

1.i *Spirit*

Spirited, of animals

(adjectives) proud c1407-1780 *chiefly poet*

In a spirited manner

(adverbs) jollily 13..-1581; heartily c1330—

To show spirit

(phrases) To do muster 1436 *obs*

1.iii *High spirit*

High spirits

(nouns) *heort*/heart OE—; fierceness c1400-c1489+1692 *obs*; animosity 1432/50-a1670 *obs*

High spirited

(adjectives) *modig*/moody OE-a1250+1375+a1400/50+1606-1755 *dial*; fierce 1297-a1533 *obs*; jolly c1330-1642; heartly 1340-1535; fiery c1385—.

In a high spirited manner

(adverb) heartly a1300-c1450

1.iv *Keenness*

The quality of keenness

(nouns) *mod*/mood OE-a1500+1579 *obs*; keenship c1205+a1225 *obs, rare*; feerness 1475 *obs, rare*.

One who is keen

(noun) lion c1175-1863.

(n phrase) marmaric(an) lion c1470-1607.

Keen

(adjectives) *hwaet*/what OE-c1205; *caf*/cof OE-c1330 *obs*; keenly c1205+1570 *obs, rare*; sternful c1400-a1500+1849 *arch.*

In a keen manner

(adverbs) *modelice*/moodily OE-1603 *obs; cenlice*/keenly OE-c1400 *obs*; keen c1475 *obs.*

Historical Thesaurus of English

Information about the Thesaurus

EDITORS
Christian Kay
Irené Wotherspoon
Jane Roberts
M.L. Samuels

INSTITUTE
Department of English Language, University of Glasgow

ADDRESS
Enquiries to:
Professor C. Kay
Department of English Language
University of Glasgow
Glasgow G12 8 QQ
UK

e-mail: CJKay@human.gla.ac.uk
tel.: +44 (141) 330-4150
fax: +44 (141) 330-3531

web site: http://www.arts.gla.ac.uk/EngLang/thes.htm

Contents of the Dictionary

TYPE

The *Historical Thesaurus of English (HTE)* is a conceptually and chronologically ordered thesaurus of meanings in the English language. The estimated 650,000 entries are arranged under taxonomic semantic headings constituting short definitions and subdivided by part of speech. There are three major semantic groupings, The External World, The Mind, and Society, followed by 26 main category headings, which are in turn divided into semantic fields and subfields. Entries are given in modern English spellings, and are accompanied by dates of recorded use and by regional and stylistic labels where appropriate.

CORPUS

The primary corpus is the *New Oxford English Dictionary* (*NOED*: 1989), which includes the integrated Supplements. Senses of polysemous words are recorded separately and treated under as many headings as are required by the meanings. Coverage aims to be comprehensive with the exception of certain closed class words, nonce-words, transparent compounds of limited attestation, words restricted to later dialect use, and specialist levels of modern scientific taxonomies. However, the editors have discretion to include such words if they think they are of interest. Place and personal names occur mainly in attributive use, as in *Stilton cheese* in the Food category. Verbs are given in the infinitive form without *to*.

Because *NOED* records mainly Old English words which survived into Middle English, the Old English materials have been supplemented from the corpus of the *Thesaurus of Old English* (see separate article). Words recorded after the publication of *NOED* are being supplied by Oxford University Press and added to the corpus.

DESCRIBED PERIOD

Coverage ranges from the earliest recorded Old English in the eighth century AD to the modern English of the 1990s. Dates are those given in *NOED*, even where these are disputed, since consistency of dating is thought to be important and it is assumed

that dating in earlier periods is always going to be approximate. Old English words are dated simply as "OE".

DESCRIBED AREA

The main focus is on the development of English within the British Isles. Regional forms within Britain are labelled using the *NOED* system of labels. Other major varieties of English, such as American, Australian or Caribbean English are included as they occur in *NOED*, suggesting a degree of British usage. There is no reason in principle why material from other historical dictionaries could not be added, but that lies beyond the scope of the present project.

EXAMPLE OF AN ENTRY

Starting with the broad tripartite division, the lexical entries are classified in numbered semantic categories of decreasing levels of generality. Within each category, words which bear a loose relationship of synonymy to the heading are listed. A system of dots is used to represent relationships of subordination and co-ordination among subcategories. At each stage in the taxonomy, words are arranged in chronological order, thus allowing access to the terms used for particular concepts in particular periods. Each word is followed by its first date of use as recorded in *NOED*, and, as appropriate, its last date, or a dash to indicate currency up to the present, or a "+" sign to indicate interrupted currency, and any restrictive labels. Main category headings begin with capital letters and subordinate headings with lower case.

History of the Project and Planning

SHORT HISTORY OF THE PROJECT

The project was begun in 1965 by M.L. Samuels, Professor of English Language at the University of Glasgow, in co-operation with departmental colleagues. The initial stage involved transferring material from the *Oxford English Dictionary (OED)* to paper slips, which were then filed under the 990 conceptual headings of Roget's *Thesaurus of English Words and Phrases*. Jane Roberts, now

Professor of English at King's College, London, undertook the collection of supplementary Old English materials (see the report on *A Thesaurus of Old English*, elsewhere in this book). In 1969, a grant was received from the Leverhulme Trust to employ two research assistants, Christian Kay and Irené Wotherspoon, who now edit the project. Substantial support has subsequently been received from the Leverhulme Trust, the British Academy and the University of Glasgow. In addition to the work of resident research and computing staff, contributions to the databank and to classification have been made by academic colleagues and postgraduate students as spin-offs from their research. Much of the data entry has been done by trainees on government training programmes.

In the 1970s a decision was made to incorporate material from the *OED* Supplements as they were published. A sampling exercise produced the broad outlines of the semantic classification and work began on classifying the materials according to this system. The basic principle, however, is that, at all but the top levels, the classification should emerge from the data rather than be imposed upon it. In the 1980s, the basic archive was completed and the decision made to enter the materials into an electronic database. The concentration since then has been on refining and completing the classification.

PLANNED YEAR OF PUBLICATION

This is difficult to estimate, since it depends on continuing success in raising funds for salaries. If all goes well, we expect to have classification and proof-reading completed by 2000 and to publish soon thereafter. In the meantime, subsets of data can be made available to accredited scholars.

FORM OF PUBLICATION

HTE will be published first in print (Oxford University Press) and then electronically, either on CD-ROM or on the Internet. Ideally, the electronic version will run in conjunction with *NOED* so that users can work between them. The printed version is expected to occupy two large volumes. It will have a detailed explanatory introduction and an alphabetical index. While each heading will be numbered in the database, we do not envisage using the full depth of

numbers in the printed work, where subordination will be indicated by typography.

Lexicographical Tools and Methods

The data were originally recorded on paper slips, and these remain the basic unit of operation, used by the classifiers to sort and number the entries and then transferred to disk in classified order for storage, editing and eventual printing.

THE HARDWARE

We have used a variety of PCs over the years; standard PC 486s with 8MB RAM and 250 MB HD are currently the basic data entry machines. These are linked in a network so that each day's output can be backed up, checked and transferred to storage on the local Unix server which has replaced the mainframe computer. *NOED* on CD-ROM is an essential editing tool.

THE SOFTWARE

The operating system of the PCs is MS DOS 6.2, and the LAN uses MS Windows for Workgroups.

The development of the database has passed through four phases:

Phase 1: dBase 11.4. This was chosen because of its compatibility with printing requirements. dBase is still used as an initial data-entry program, but did not prove powerful enough for searching the fast-growing corpus on PC in an acceptable time.

Phase 2: CAFS. (Computer Addressable File-Store). This was used for an ICL mainframe version because it was the only powerful search engine then available. It was abandoned when the University withdrew mainframe facilities.

Phase 3: Ingres. The Ingres database currently in use was designed by Irené Wotherspoon to achieve maximum speed and flexibility in retrieving data. It runs on a Unix server. It

offers the advantage of compatibility with other university projects, especially the thesaurus projects at King's College, London, with whom we work in close co-operation.

Phase 4: Access. Flora Edmonds is currently working on developing a user-friendly front-end in Access.

The database consists of 29 fields, covering the word (or heading), part of speech, dates, style labels, the numerical codes as illustrated, a further number which makes each word uniquely retrievable, Roget numbers, and cross-reference numbers under our system of classification. Our general policy is to incorporate common search questions into a menu system while making the full power of the database available to users who are prepared to learn the query language.

Special Remarks

PURPOSE OF THE PROJECT

The primary purpose of *HTE* is to provide a research tool for work on the history of the English Language. Our theoretical starting point is that a more precise picture of the development of the lexicon can only emerge when we are able to study and compare such factors as obsolescence, innovation and transference within and between fields of meaning.

Studies of individual fields in these terms have already been undertaken in several dissertations (Chase 1983, Coleman 1992, Ower 1996, Reay 1991, Sylvester 1991, Thornton 1988, Wotherspoon 1969). These studies also involved their writers in consideration of factors relevant to particular fields, where connections are made with sociolinguistics, semantics and stylistics, or, more broadly, social history, the history of ideas and the study of literature. There are also implications for semantic theory, both synchronic and diachronic.

SAMPLE ENTRY

The entry for *Gin* is given below. It should be noted that this sample may well be incomplete; any new words or words which have been redirected from other categories since this section was keyed in will be added at the final editorial stage.

The full numerical heading for *Gin* is: 01 02 08 02 02 06 03.
This represents:

01	The External World
01 02	Life
01 02 08	Food/drink, eating/drinking
01 02 08 02	Beverage
01 02 08 02 02	Intoxicating liquor
01 02 08 02 02 06	Distilled drink
01 02 08 02 02 06 03	Gin

The list of exponents of the heading is:

diddle c1700-1858 *sl*; geneva 1706--; gin 1714--; strike-fire 1725 *sl*; tittery 1725-1751 *sl*; royal bob 1729-1770; strip-me-naked 1751 *sl*; sky-blue 1755-1796 *sl*; lightning 1781-1851 *sl*; jacky 1799-1832 *sl*; max 1811-1851; daffy 1821-sl; ribbon 1823-*ca*; sweet-stuff 1835; bottled lightning 1839; (white)satin 1845-- *sl*; tiger's milk 1850 *sl*; juniper 1857 *sl*; cream of the valley 1858; eye-water 1869-- *sl*; cream of the wilderness 1873; twankay 1900 *sl*; mother's ruin 1937-- *sl*; panther 1942-- *sl*.

ca = cant; *sl* = slang

The Historical Thesaurus of English
Summary of Classification

Section I:

THE EXTERNAL WORLD

(largely complete)
1. The Earth
2. Life
3. Sensation & Perception
4. Matter
5. Existence
6. Relative Properties
7. The Supernatural

Section II:

THE MIND

(in progress)
1. Mental Processes
2. Emotion
3. Good or Bad Opinion
4. Aesthetic Opinion
5. Will
6. Endeavour
7. Language
8. Possession

Section III:

SOCIETY (largely complete)
1. Social Groups
2. Habitation
3. War
4. Government
5. Law
6. Education
7. Religion
8. Communication
9. Travel and Transport
10. Work
11. Leisure

Department of English Language
University of Glasgow
Revised 1995

Dictionary of Old English

Information about the Dictionary

EDITORS

Chief Editor: Antonette diPaolo Healey
Revising Editor: Joan Holland
Editors: David McDougall
Ian McDougall
Nancy Speirs
Pauline Thompson
Editorial Staff: Catherine Georgi (copy-editor)
Elaine Quanz (computer-editor)
Takamichi Ariga (systems analyst)

INSTITUTE

Centre for Medieval Studies, University of Toronto

ADDRESS

Correspondence is to be sent to:
Dictionary of Old English
Room 14285, Robarts Library
University of Toronto
Toronto
Ontario M5S 3H1
Canada
e-mail: Healey@doe.utoronto.ca
tel.: +1 (416) 978-8883
fax: +(416) 978-8835

Contents of the Dictionary

TYPE

The *Dictionary of Old English* (*DOE*) is an historical dictionary, providing a comprehensive record of the earliest period of the English language. The c. 35,000 entries of the completed dictionary will offer information on grammar (together with a full list of attested spellings for each word, arranged in parsed order); number of occurrences (and sometimes patterns of usage); definitions (with supporting citations); Latin equivalents, if any; related Old English words; and finally, Middle English and Modern English reflexes, if any. The spelling of the headword is late West Saxon if attested in that dialect; if not, the most frequently attested spelling is chosen as the headword spelling. Nouns and pronouns are given in the nominative singular, verbs in the infinitive.

SUBJECT/CORPUS

The dictionary is based on a closed corpus of Old English texts originally written between 600 and 1150 AD, although some material is extant today only in transcripts as late as the eighteenth century. The highest proportion of the Old English material is from the eleventh century. Approximately 3,000,000 running words of Old English will eventually be ordered into c. 35,000 entries. The texts of the Corpus, mainly edited in the nineteenth and twentieth centuries, fall into five broad categories: prose (69%), Old English glosses to Latin texts (24%), poetry (6%), glossaries (1%), and inscriptions (under 1%). The dictionary excludes three types of information: etymologies, place names, and personal names. They have been excluded either because a separate dictionary of the appropriate type is being compiled or because the research tools needed to produce adequate entries are not yet available. Once this material is published, it can be integrated into the entries of the *Dictionary of Old English.*

DESCRIBED PERIOD

The period the dictionary surveys begins with the reign of Æthelberht of Kent (597-616) and the code of laws issued during his reign to the closing annal of the *Anglo-Saxon Chronicle* in 1154

when Henry II succeeded Stephen to the throne of England. However, we are trying to ensure that no material of the transitional period between Old and Middle English is overlooked, and occasionally we have extended the upper limit of our period in order to encompass late twelfth- and thirteenth-century texts which show the development of the earlier tradition. Unlike the practice of the *MED* and the *OED*, we do not give the date of composition of our material because for the Old English period this is often a matter of great uncertainty.

DESCRIBED AREA

Anglo-Saxon England is the term applied to those areas of Great Britain where the vernacular Germanic language, Old English, was written. It encompassed most of what now is England and the lowlands of Scotland, south of the Firth of Forth and of the Clyde. It did not include what now forms Wales, northern and south-western Scotland, and parts of the south-western peninsula. Old English is also preserved in continental manuscripts.

EXAMPLE OF AN ENTRY

Each entry in the *DOE* can provide up to twelve fields of information:

1) the headword, normally in its late West-Saxon spelling; 2) cross-references for every main vowel or consonant, and for any other spelling which might be puzzling; initially stored with the entry, these cross-references appear in the fascicle sorted alphabetically among the headwords and other cross-references; 3) grammatical information; 4) a full list of spellings, arranged in parsed order; spellings unique to the thirteenth century and later are given separate treatment and labelled 'late'; 5) number of occurrences and usage labels if significant (e.g., restrictions in occurrence or use by date of manuscript or dialect, by genre or text or author); 6) a schema for an overview of the senses of a large word; 7) definitions, which may be analytical, contextual, or a translation equivalent, dependent upon the nature of the word; 8) citations which document the meaning(s) of the word being defined; it is our policy to cite all instances of a word with twelve or fewer occurrences; 9) collocations (e.g., *hat and ceald* 'hot and cold'); 10) Latin

equivalents, i.e., Latin words which correspond to the Old English headword and are present in the same context in the same manuscript; 11) related Old English words, which provide an implicit analysis of the word in relation to its word family; 12) secondary references, the most important of which are the Middle English and Modern English developments of the Old English headword.

History of the Project and Planning

SHORT HISTORY OF THE PROJECT

Angus Cameron, the founding editor (1970-83) of the *Dictionary of Old English,* planned for three phases of the project: the development of the research collection, computer processing of Old English texts, and finally, the actual entry-writing. The first two phases are essentially complete; we are now in the third and most exacting phase.

The *Dictionary* is written by making use of the *Dictionary of Old English Corpus in Electronic Form,* a database developed in 1981 by the project and comprised of at least one copy of every Old English text (and sometimes more than one copy if of importance for dialect, date, etc.), on which a machine-concordance has been run. This Electronic Corpus forms the citation base of the *Dictionary* and has undergone corrections as errors are found, and improvements as better scholarly editions are published and incorporated. The most recent version of the Corpus (1995) is among the first databases of a research project in the humanities to be fully conformant with the 1994 guidelines issued by the Text Encoding Initiative (TEI) P3, the current standard for encoding electronic texts. An International Advisory Committee of four scholars, a number of specialist readers in Old English language and literature, and several consultants in computing in the humanities advise the editorial team.

(PLANNED) YEAR OF PUBLICATION

Since 1986, the project has published six of its 22 fascicles on microfiche: *A* (1994), *Æ* (1992), *B* (1991), *C* (1988), *D* (1986), and

E (1996). We plan to publish the seventh letter, *F*, together with the previous six, on CD-ROM in 1999. Depending on stability of funding and editorial staff, we hope to complete the *Dictionary* by around 2013.

FORM OF PUBLICATION

The full *Dictionary* will eventually be printed on paper and published by the University of Toronto Press. It will also simultaneously be published electronically. However, since completion of the *Dictionary* will be at least 16 years in the future, it is impossible to make specific projections about the electronic means of final publication, dependent as they are on the state of computer technology.

Lexicographical Tools and Methods

THE HARDWARE

The project's first generation computer system was a hybrid system consisting of a Xerox 1108 workstation, a Xerox 8045 laser printer, and a VAX (the 11/730), all connected on an Ethernet. Our second generation system was an enhancement of the first by the addition of two 300 MB fileservers and three 1108 workstations. In 1990 we began our migration from Xerox to Sun Microsystems, moving from a closed system to an open UNIX environment and a more standard platform. Moreover, the Suns allow us to run without modification the large number of programs and systems we developed for the Xerox workstations. Our present configuration is a network of Sun Microsystems workstations (3 SPARCstation 1s, 1 SPARCstation 2, 1 SPARCstation 4) running SunOS 4.1.x, and a QMS PS-1700 PostScript network laser printer.

THE SOFTWARE

Much of the software for the project has been developed in-house. There are two principal stages in the writing of entries for the Dictionary, each with different software requirements: lemmatisation and actual entry-writing. The Lemmatisation of the Concorded Corpus is a menu-driven program, written in Lisp, which

allows an editor to assign individual spellings, or groups of spellings, from the Electronic Corpus, to Dictionary headwords. The resulting data is in the form of triplets: the headword, the actual spelling, and the citation reference. These lemmatised concordance files provide the basic data for entry-writing.

To these lemmatised files are added information drawn from our Catalogue Database. This relational database gives useful information about the more than 3000 texts which constitute the Old English Corpus, such as manuscript date, variant manuscripts, the short title of the Latin source, etc. The Catalogue Database application runs on a SPARCstation under UNIX, uses the Empress relational database management system, and is written in Empress application command language called MBuilder. The merging of the lemmatised concordance files with information from the Catalogue Database results in the production of paper slips which convey information efficiently for entry-writing.

Entry-writing is a combination of manual and electronic procedures. Paper slips are used for the actual ordering of the citations into senses. We have not yet been able to devise an electronic system that allows for the manipulation of sometimes hundreds of citations into appropriate sense divisions which is more efficient than the human hand. Once these slips are manually ordered, their identifiers are entered at the entry-input stage and the actual citations are retrieved from the Electronic Corpus.

During entry-writing, editors can search the Corpus using two tools. The TAL (Text Analysis Language) Browser, developed in-house by a student research assistant, allows searches on various indexes for spellings or addresses. It also permits both Boolean and regular expression searches. Our most recent and most powerful tool is the WEB Corpus, the interface of which was designed by Mr John Price-Wilkin of the University of Michigan, which uses the Open Text Software Suite which includes PAT, the highly sophisticated search engine developed by the University of Waterloo for searching the electronic *OED*.

Also during entry-writing, editors can search previous fascicles and any entries in progress. FascSearch is a simple menu-driven tool, written in Lisp, used to find a pattern in any fascicle or entry. Similarly, FindHW (Headword search) allows searches

through headwords to find all the headwords that have been completed to date. The information we are able to retrieve has been useful in allowing us to check our conventions.

Finally, at the production end, we have developed a Citation Checking program and automatic procedures for the pre-publication of the fascicles.

SAMPLE ENTRY

ēag-sealf — headword

Noun, f., cl. 2 — grammatical information

Att. sp.: eagsealf; eahsealf | eahsalf (Med 5.8) || eagsealfe, eahsealfe, eahsealve; heahsealuæ | eahsalfe | ehsealfe, ehsealue | ægsealfe || eagsealfa | egsealfe. — full list of attested spellings, arranged in parsed order

Late: eahsylfe (Med 1.1 MS O, xiii) — late spelling given separate treatment

30 occ. (mainly in medical texts) — number of occurrences and usage note

eye-salve, ointment for the eye — definition

Lch II (1 Head) 2.1: læcedomas wið eallum tiedernessum eagena ... & wiþ wyrmum on eagum & **eagsealfa** ælces cynnes. — block of quotations; each quotation has a short title (in bold) and a system of reference

Med 1.1 3.1: wiþ nædran slite, heortes horn hafað mægen ælcne wætan to adrigenne; for þam his man bruceþ on **eagsealfe** (H *ægsealfe,* O *eahsylfe;* cf. L: *et ideo ex eo in colliriis ocularibus utuntur*). — readings from variant MSS (H, O); Latin source material

Lch II (1) 2.14.2: wið flie **eagsealf**, genim bromes ahsan & bollan fulne hates wines.

Med 3 5.1: **eahsealf**, win & piper, do in horn, & in þa eagan þonne þu ðe restan wille.

Med 5.4.1 16: sy gemenged togædere hunig tear & win & rudan seaw & efenfela gedon on cyperen fæt oþþe mæstling oþþe bræsen, nim þær gode **eahsealfe**.

Med 5.8 1.1: ðis is seo seleste **eahsalf** wið ehwærce & wið miste ... & wið teorendum eagum.

Med 3 10.1: þis is seo æðeleste **eahsealf**, wið eahwyrce & wið miste, ... & wið tyrendum eagan.

Lch II (3) 2.6.8: wyrc smeþe **eagsealfe** nim buteran, wyl on pannan, afleot þæt fam of & ahlyttre þa buteran on blede.

AldV 13.1 3051: *malagma .i. colirium* **ehsealfe**, þone halwendan cleoþan (AldV 1 2950 *ehsealfe*, AldV 9 224 *heahsealuæ*, AldV 10 141 *ehsealue*, AldV 14 137 *eahsealve*; from ALDH. Pros.virg. 34, 276.2 *visum scilicet caecis et malagma monoptalmis impertiendo*).

AntGl 2 336: colliria **egsealfe** (Jun.Transcr. 71 *eagsealve*; perh. from ISID. Etym. 4.9.10 *collyria latinum sonat, quod vitia oculorum detergant*). — reading from a late transcript (Jun.Transcr.)

Lat. equiv. in MS: collyrium, malagma — Latin equivalents in the same MS

See also: ēage, sealf — related Old English words

MED ei(e-salve. OED2 eyesalve. — secondary references

Middle English Dictionary

Information about the Dictionary

STAFF

Many people have contributed to the editing and production of the dictionary since its beginnings in 1930, including former editors Samuel Moore (1930-34), Thomas A. Knott (1935-45), Hans Kurath (1946-61), and Sherman M. Kuhn (1961-82/83). The current staff consists of:

Editor-in-Chief:	Robert E. Lewis
Review Editors:	Elizabeth S. Girsch, Marilyn S. Miller, Mary Jane Williams
Associate Editors:	Mona L. Logarbo, Paul F. Schaffner, Jeffrey L. Singman
Production Staff:	Lidie M. Howes, Head; Patrick M. Florance, Elaine L. Halleck, William R. Hosticka, Ekaterina Vladimirsky

ADDRESS

Correspondence should be sent to the editor-in-chief at:
Middle English Dictionary
University of Michigan
555 South Forest
Ann Arbor, Michigan 48109-2531
USA

e-mail: RELewis@umich.edu
tel.: +1 (313) 998-7110

Contents of the Dictionary

TYPE

The *Middle English Dictionary* (*MED*) is an alphabetically ordered, historical period-dictionary. Each entry begins with various kinds of information (see EXAMPLE OF AN ENTRY, below), followed by the meanings illustrated by chronologically arranged quotations. The spelling of the entries is, in principle, that of the Southeast Midland dialect of Middle English. Nouns are usually given in the nominative singular form (though occasionally in the plural if only attested in that form). If the evidence is extensive enough, pronouns are given separate entries for nominative, accusative, genitive, and occasionally dative cases, and the same applies to the definite articles in *th-* derived from Old English. Verbs are given in the infinitive. Participles are treated under the verb unless they occur only in adjectival constructions, in which case they are given separate entries. Compounds and combinations are treated under one or more of the elements involved unless there is some complexity of meaning, form, or etymology, in which case they are given separate entries.

SUBJECT CORPUS

The dictionary is based on a collection of quotations from Middle English texts assembled over a number of years, beginning with donations in 1930 from the Oxford English Dictionary (*OED*) of all of its Middle English slips and from Cornell University of the Chaucer slips it inherited from Ewald Flügel along with the supplementary slips collected there under the auspices of the Modern Language Association of America between 1925 and 1930. The further reading and extracting of quotations from Middle English texts was one of the primary activities of the dictionary staff, helped by a number of volunteers, between 1930 and 1945, and this has continued to the present for newly edited or re-edited texts. The collection now contains some three million quotations from all genres of Middle English (including surnames and place names), taken primarily from printed editions but supplemented for works and genres not available in printed editions (documentary material, medicine, popular science, etc). from manuscript facsimiles.

DESCRIBED PERIOD

The period covered by the dictionary is 1100 to approximately 1500. Inevitably there are difficulties in breaking up the continuum of any language into discrete periods, and this is no less true of the Middle English period than of other periods: there is some slight overlap with Old English at the beginning and some more pervasive overlap with Early Modern English at the end. At the beginning there are, for example, some texts from the twelfth and thirteenth centuries that are copies of Old English originals; they have been treated as Middle English, based on their date of production and on their slight variation from the Old English, but they could just as easily, based on their language, be called late Old English, and some of them are also included in the *Dictionary of Old English* (in progress; see elsewhere in this book). At the end the normal cut-off date for manuscripts is 1500, but books printed in the fifteenth century are excluded, as are documents (as distinct from literary texts) written later than the early 1480s, even if preserved in manuscripts, on the grounds that the invention of printing standardised spelling and that documentary writing became standardised and conventionalised in the second half of the fifteenth century. These are arbitrary decisions, and no real harm is done if transition texts at either end of the continuum are counted in two periods; the greater harm would be the possibility of words and forms from these transition texts being omitted in the relevant dictionaries.

DESCRIBED AREA

All Middle English texts written in England, Ireland, and Wales are included in the corpus for the dictionary; Scottish texts are excluded because they are being treated by the *Dictionary of the Older Scottish Tongue* (in progress).

EXAMPLE OF AN ENTRY

Each entry in the *MED* begins with the headword followed by the part of speech; then the variant spellings and forms that are attested in the quotations, with labels for those that are clearly dialectal or early (1100 to 1300); cross-references, where appropriate, to other entries; and the etymology. (Reconstructed diacritics are given, where appropriate, on all headwords, on certain of the following variant

spellings and forms, and on all forms in the etymologies). The body of the entry contains the definitions, arranged in logical order, followed by illustrative quotations for each one, beginning with the earliest and concluding with the latest, with at least one quotation for each quarter-century in between if available. For an example of an entry see SAMPLE ENTRY, below, and for a full treatment of how an *MED* entry is put together, see Sherman M. Kuhn, "On the Making of the *Middle English Dictionary*", *Dictionaries: Journal of the Dictionary Society of North America* 4 (1982), 14-41, especially 26-39.

History of the Project and Publication

EARLY HISTORY OF THE PROJECT (1930 TO 1952)

The idea of a dictionary of Middle English on historical principles goes back to W.A. Craigie, the third editor of the *OED*, who in 1919 proposed a series of national and period dictionaries, which began to take shape in the 1920s. The present *MED* began life at the University of Michigan in 1930 with three events: the donation by the *OED* of its collection of Middle English slips (along with its collection of Early Modern English slips for the Early Modern English Dictionary begun by C.C. Fries a few years earlier); the transfer from Cornell University of Ewald Flügel's Chaucer slips and the Middle English supplementary material collected there between 1925 and 1930; and the appointment of Samuel Moore from the University's Department of English as editor. During Moore's editorship (1930-34) the collection of quotations was tested for inclusiveness and accuracy, an extensive supplementary reading program was planned and carried out, a dialect survey was completed, and a general editing plan was formulated in consultation with scholars in the United States and Europe; the dialect survey was subsequently published as "Middle English Dialect Characteristics and Dialect Boundaries" by Moore, Sanford B. Meech, and Harold Whitehall in *Essays and Studies in English and Comparative Literature* (Ann Arbor, 1935), pp. 1-60. During Thomas A. Knott's editorship (1935-45) the reading program was extended into areas that had been slighted before (Latin and Old French documents

containing English names for household utensils, tools, local taxes, etc., scientific and technical works, and place and personal names), but the serious economic disturbances of the depression and a reduction of staff during World War II slowed overall progress. During the first few years (1946-52) of Hans Kurath's editorship a formal editing plan was drawn up, and editing began on the letter E and then progressed to F (A to D had already been edited according to an earlier plan but were postponed for re-editing until F had been completed according to the new plan). During these early years the project was supported almost exclusively by the University of Michigan, with some assistance in the 1930s from the American Council of Learned Societies and the Rockefeller Foundation.

PUBLICATION AND LATER HISTORY (1952 TO THE PRESENT)

The first fascicle (the early part of E) was published by the University of Michigan Press in 1952, and between then and 1984 fascicles appeared at an average rate of two per year, first the letters E and F, then A to D and G to P. In 1954 a *Plan and Bibliography* was published, describing the editing plan and methods and containing the list of short titles used for the citations, with the relevant bibliographical information. Fascicles are customarily 128 pages, and up to 1984 the camera-ready copy for them was produced on two IBM Executive typewriters (one for the secretarial font and the other for the boldface font). Since 1984, with the beginning of Q and a supplement to the *Plan and Bibliography*, we have been using a computer-assisted production system, which has allowed us to update and revise our fonts and to increase our rate of publication to an average of three fascicles per year and (in early 1997) to reach the end of T. The editing should be completed by the end of 1997 and the reviewing and the production of the last fascicles by the middle of 2001. By that time, with the planned keyboarding of A to P, the mark-up of Q to T currently in progress, and the use of SGML for entry beginning with U, the dictionary will be ready for online publication in addition to its traditional print publication. The project continues to be supported by the University of Michigan, supplemented by grants from the Andrew W. Mellon Foundation

(since 1975) and the National Endowment for the Humanities (since 1980) and by contributions from private donors.

Special Features

SHORT TITLES AND DATING

One of the innovative features of the *MED*, which derives from the early years of Kurath's editorship, is its treatment of manuscripts and texts. Because the purpose of the dictionary is to describe, or recover, the English language current between 1100 and approximately 1500, we are interested in what the actual witnesses had to say, that is, what the medieval scribes put down on parchment and paper, and our bias toward these witnesses can be seen in our basic, overall approach to our material. First of all, "For [nearly] every text and for [nearly] every version of a text a preferred MS is selected, from which the quotations are normally taken .... The earliest complete, or fairly complete, MS that has been edited is chosen as the preferred MS" (1954 *Plan and Bibliography*, p. 17). This usually turns out to coincide with what the editor of the text believes is the most important manuscript or the one closest to the author's original, though "If the most important MS ... has not been edited, or poorly edited, a facsimile of that MS may be used as the preferred MS" (p. 17).

The manuscript date is the first item in the short title for nearly every text and takes precedence over the composition date, which is usually conjectural (see the following SAMPLE ENTRY for examples). If the composition date is more than 25 years earlier than the manuscript date, then we put it in parentheses following the manuscript date, though only in the short title for the preferred manuscript; otherwise we omit it, leaving just the manuscript date. There are some exceptions, or rather refinements, to this practice, but the principle still holds: the manuscript date is the one we consider the more important; and indeed we are constantly revising our manuscript datings as new evidence appears.

"For each MS of a text that has been edited by two or more different scholars, a preferred edition is selected, from which the quotations are normally taken .... In selecting the preferred edition,

fidelity in the reproduction of the MS is the primary consideration" (p. 17); in the case of competing critical editions, the ease with which we can recover and restore the manuscript readings is the crucial factor in our choice of preferred edition. In short, we treat each manuscript, whether preferred or non-preferred, as a witness, and our short title stands for that witness – not for the edition in which it happens to appear. If we believe that a manuscript, either preferred or non-preferred, reproduces a garbled or erroneous reading, we sometimes add a bracketed reconstructed reading after it to indicate what we think the reading was intended to be or was derived from, but we leave the manuscript reading intact. We thus preserve the integrity of the scribe, the actual witness, and of his manuscript, his testimony.

ETYMOLOGICAL PRACTICE

A number of dictionaries of English have very good etymological notes, including the *OED* itself as well as more recent unabridged dictionaries like *Webster's Third New International Dictionary of the English Language* (1961) and *The Random House Dictionary of the English Language*, (2nd edn, 1987), and specifically etymological dictionaries like the *Oxford Dictionary of English Etymology* (1966) and the recent *Barnhart Dictionary of Etymology* (1988), not to mention some of the dictionaries of the languages from which Middle English words are most often derived. It has seemed unnecessary to repeat or paraphrase in the *MED* the elaborate etymologies in such dictionaries, which trace words back through Germanic, Italic, etc. to their ultimate Indo-European sources. Our practice, as formulated in the 1954 *Plan and Bibliography*, is that "Only the immediate source of the M[iddle] E[nglish] word is given .... The ancestor of the ME word is cited in the form (or forms) that will enable the reader ... to pursue its remoter history in the standard dictionaries of O[ld] E[nglish], O[ld] F[rench], M[edieval] L[atin], O[ld] I[celandic], etc. Variants of the source language that appear to be reflected in the ME variants or spellings are cited ... along with the 'standard' form" (p. 7).

The Middle English period is the great formative period in the development of the English lexicon, and the reason for this is that Middle English borrows or assimilates from so many different

sources: Old English; Old Norse, primarily in the northern and eastern parts of the country; French, both in its Anglo-French form early in the period and in its Central French and literary forms later in the period; Latin, both in its classical literary form and in its medieval form; and the continental Germanic dialects of the Low Countries, Middle Dutch and Middle Low German; and the characteristics of, and the relationships among, these five languages are often complex. For these reasons it is important not just to give the standard headwords in the dictionaries of the source languages, but, because we are dealing with languages in contact, competing side by side in England, to illustrate the actual forms in these languages that might have given rise to the Middle English forms, including sometimes even the errors. During Kurath's editorship the etymologies were generally terse, but during Sherman Kuhn's editorship (1961-82/83) they became fuller and more elaborate, especially for those from Old English, and since then the fullness has extended to the Latin and Old French etymologies as more Anglo-French evidence has become available and as we have looked further into the Medieval Latin and Anglo-Latin evidence.

NAMES

During Knott's editorship the reading program was extended into place and personal names, and since then we have drawn on the data in the volumes of the English Place Name Society and the various collections of surnames that have been published with increasing frequency since the 1930s. We do not try to be complete in our use of the name-evidence, however; we use it only as corroboration for words attested in other texts, and one of the primary reasons for that use is etymological. "When, for example, a word appears in O[ld] E[nglish] and then disappears from ordinary written sources to turn up once more in the fourteenth or fifteenth century, one may be in doubt as to whether the later instances were really descended from OE or whether they represent a new formation in M[iddle] E[nglish] or borrowing from another language. But if the word is well-attested in names of the twelfth and thirteenth centuries, one can accept the OE etymology with greater confidence" (Kuhn, pp. 20-21). Similarly, name-evidence frequently gives us our earliest occurrences of words borrowed from Anglo-French and Old Norse. With such supplement-

ary evidence we are able to corroborate etymologies arrived at through the standard dictionaries as well as to demonstrate closer chronological links to the various source languages and dialects in question, and thus to provide the fullest picture of what we call the "immediate" etyma.

SAMPLE ENTRY

trusse n. Also **trus(e, trous, tros(se;** pl. **trusses,** (early) **trussen, trossen.** [OF **tros(se, trous(e,** AF **trus(s)e** (vars. of OF **torse**) & ML **trossa, trousa,** AL **trussa.**]

Headword and part of speech; variant spellings; etymology

1. (a) A wrapped bundle containing clothing, laundry, etc., a parcel, traveler's pack; also, ?a bolt of cloth, cloth wound and fastened together in a roll [quots. c1436]; **in ~,** ?on the bolt; (b) a bound sheaf of harvested grains, etc.; a bundle of sticks, firewood, etc.; also, a quiver of arrows; *(c) mil.* a wad, pad, or roll of heavy material hung on a fortification to defend against a battering ram.

First sense division, with subdivisions

Subdivision with field label (=military usage)

(a) **?c1225** *Ancr.*(Cleo) 129/8: Noble men & gentile ne beoreð nane packes ne ne fareð wið trusses ne wið purses itrussed. **c1230(?a1200)** *Ancr.*87/23: Trussen & purses, baggen & packes beoð worltliche þinges. **(c1391)** Gower *CA Suppl.* (Hnt) 7.3253*: The paien rod upon an asse...With him a riche trusse he ladde. **c1436** *Ipswich Domesday(2)* 181: Of smal trussys of clothys, tunnes of woide, barell of syndres, and all other thynges oute takyn wynes, it is ordeyned that other power, porturys or bererys, moun travaylyn among these thynges. *Ibid.*187: Other clothes...ben bought in the cuntre...the which clothes be in trusse or in pakke or in fardel. **(1440)** *PParv.*504: Trusse, or fardelle: Fardellus, sarcina. **a1450** *Gener.(1)* 7040: Ech of vs a trusse shuld bere, Wonden in a feir white shete, On oure heides, that whom we mete Shuld wene that I youre seruaunt were, You to help youre clothes to bere. **a1500(a1450)** *Gener.(2)* 4425: She was so ferde hir truse was in falling. **c1650(a1450)** *Death & L.*11: When Death driueth att the doere...noe truse can be taken, noe treasure on earth.

Illustrative quotations for each subdivision, in chronological order

(b) **c1325** *Mon in þe mone* 15: Wher he were o þe feld pycchynde stake, for hope of ys þornes to dutten is doren, he mot myd is twbyl oþer trous make oþer al is dayes werk þer were yloren. **(a1393)** Gower *CA* 5.5056: His Ape...hadde gadred al aboute Of stickes hiere and there a route And leide hem redy to his hond Wherof he made his trosse and bond. **(1393)** *Plea & Mem.R.Lond.Gildh.*205: [Four bows with one] trusse [of] setes, [20 s.]. **c1400** *Femina* 28: 3oure

shoffez in sheffez byndyþ, In trussez 3our sheffez putteþ. **a1425(?a1400)** *RRose* (Htrn) 4004: Vndir his heed no pilowe was, But in the stede a trusse of gras.

(c) **(a1460)** *Vegetius(2)* 2404: Yet wil thei this delude, And with oo crafte thoo craftis iii conclude: Of quylt & felt a trusse thei depende, Ther as the ram entendeth for toffende.

2. *Naut.* (a) A rope, line, cord, etc. serving as part of a ship's rigging, prob. to hold a sail to a mast or to furl or unfurl sails; (b) in cpds. & combs.: ~ **line (rop),** a rope serving as a truss; ~ **pareil,** a rope used as a parrel; ~ **puli,** a pulley attached to a truss, prob. a device for raising and lowering a sail [see also **puli** n. (b)]; **fore ~,** a rope attached to the foreyard; **main ~ (pareil,** a rope or parrel attached to the mainsail [cp. **main** adj. **3.**]; **mesan ~,** a rope attached to the mizzen sail.

(a) **(1295)** *Acc.Shipbuilding* in *Ant.J.7* 434: Item, in xii globis fili ad duas cordis que dicuntur Trossen inde faciendas. **(1296)** *Newcastle Galley* in *Archaeol.Ael.4.2* 188: Vnum par de Trusses Inuentis in domo Johannis de Pytingtone. **(1336–7)** in Sandahl *ME Sea Terms 2* 117: In diuersis cordis de Russhewale cum schiuis et Trussis pro vno rakke inde faciendo. **(1356)** *in Pipe Roll 32 Edw.III m.33/2* [OD col.]: Computat in...iiij trusses, iiij poleankeropes. **(1413)** *Inquis.Miscel.(PRO)* 7.248: [2 pairs of] bowlynys, [2] tregettys, [a couple of] trussys. **(1420)** **Foreign Acc.3 Hen.6 F/2 dorso* [OD col.]: j corda vocata slyngerope, j hauser pro Prialle ropes, j hauser pro trusses.

(b) **(1301–2)** in Sandahl *ME Sea Terms 3* 108: Trossropes, ij steteropes et ij stetingropes. **(1336–7)** in Sandahl *ME Sea Terms 2* 123: Pro vno wyndyngrop...pro duobus Trusseropes ...pro vno Botrop...[etc.]. **(1364–5)** *Ibid.:* ij bowelyns, ij trusserops. **(1402)** *Acc.Exch.K.R.43/6 m.2 dorso* [OD col.]: In Bafyngera...vocata la Gode Grace...De poleys ix, De Truspoleys ij. **(1405–10)** in Sandahl *ME Sea Terms 2* 121: ij Stetynges debil', ij Truslynes debil'. **(1409–11)** *Ibid.*122: Vn Sailyerd fait de deux peces, vn Trusparaill' pour mesme le mast. **(1420)** **Foreign Acc.3 Hen.6 K/2* [OD col.]: j trusse parraille. **(1485)** in Sandahl *ME Sea Terms 2* 38: For trusses, ij; ffor yerdes j; ffortakkes, ii; fforshetes, ij; fforlyftes, ij. *Ibid.*71: Mayne trusses. *Ibid.*80: Meson trusses. **(1497)**

Second sense division, with field label (=nautical usage) for all subdivisions; subdivision containing compounds and combinations

Cross-references to other entries in the dictionary

*Ibid.*71: Mayne truse parell ffeble.

3. Glossing AL **ansea** [?read: **ansa**]: ?a coupling for attaching curtains, a curtain ring [cp. sense **2.**(b)]; ?an architectural brace, a truss [some other word may be intended; the sense of the L word is not established].

Third sense-division

?a1500 *Trin-C.LEDict.Suppl.* 564/16: Ansea: a trusse.

4. In surnames.

Names are always the last sense-division in an entry

(1202) in Reaney *Dict.Br.Surnames* 355: John Trusse. **(1211–12)** *Red Bk.Exch.* in *RS 99* 602: Ricardus Trus. **(1327)** *Nickname* in *SAU 63* 196: Joh. Trusse. **(1327)** *Sub.R.Sus.*in *Sus.RS 10* 129: Johne Trusseman.

Mittelhochdeutsches Wörterbuch 1050-1350

Middle High German Dictionary
1050-1350

Information about the Dictionary

EDITORS

Editors-in-Chief:	Kurt Gärtner (Trier)
	K. Grubmüller and K. Stackmann (Göttingen)
Editors:	R. Plate, U. Recker, J. Tao (Trier)
	G. Diehl, B. Schnell (Göttingen)

INSTITUTES

1. University of Trier, Arbeitsstelle für das Mittelhochdeutsche Wörterbuch, and Kommission für das Mittelhochdeutsche Wörterbuch of the Akademie der Wissenschaften und der Literatur in Mainz; 2. University of Göttingen, Arbeitsstelle für das Mittelhochdeutsche Wörterbuch, and Kommission für das Mittelhochdeutsche Wörterbuch of the Akademie der Wissenschaften in Göttingen.

ADDRESSES

Correspondence should be sent to:

1. Prof. Dr Kurt Gärtner
Universität Trier
FB II Sprach- und Literaturwissenschaft
D 54286 Trier
tel.: +49 (651) 201-2323/2321
e-mail: Gaertner@uni-trier.de

2. Prof. Dr Klaus Grubmüller
Universität Göttingen
Seminar für Deutsche Philologie
Humboldtallee 13
D 37073 Göttingen
+49 (551) 39-7524/7525
mhdwb@mail.sub.uni-goettingen.de

Contents of the Dictionary

TYPE

The *Mittelhochdeutsche Wörterbuch* is an alphabetically arranged, scholarly, diachronic dictionary. It will comprise c. 80,000 entries. Each entry will define the grammatical category, describe meaning and – where appropriate – give the earliest occurrence and explain inflection. Usage will be illustrated by quotation. Headwords will be normalised, nouns and pronouns will be given in the nominative singular and verbs in the infinitive.

SUBJECT/CORPUS

The dictionary is based on a machine-readable corpus taken from scholarly editions of texts of every possible genre, covering the period from 1050 to 1350. From this corpus the occurrences of each lemma will be collected into a database (*Belegarchiv*) using data processing techniques. The material for the database is derived from the corpus in the three different ways: "A" represents the core corpus where all occurrences of all words are considered, "B" contains texts from which the occurrences of high-frequency words are excluded and "C" represents texts from which only the occurrences of low-frequency vocabulary and special usage are considered, with the help of glossarial evidence.

DESCRIBED PERIOD

The dictionary covers a period of 300 years from 1050 to 1350, centred around the classical German literature of the turn of the thirteenth century, to which belong the works of Hartmann von Aue (*Erec* and *Iwein*), Wolfram von Eschenbach (*Parzival* and *Willehalm*), Gottfried von Straßburg (*Tristan*), the *Nibelungenlied* and Walther von der Vogelweide's songs. The earliest major text for the dictionary is the Commentary to the *Cantica Canticorum* by the Benedictine abbot Williram of Ebersberg from about 1060, and a major source from the end of the period is the collection of sermons by the Dominican mystic Tauler from before 1350. The criteria for placing the texts of the corpus within the given period are twofold: the first and main criterion is the date of composition of a text, and the second, if the date of composition is uncertain, is the century or

decades in which the transmission was greatest according to the quality and number of manuscripts still existing.

DESCRIBED AREA

The geographical area of Middle High German covers the whole of the High German area south of the High German sound-shift isogloss. This area comprises the Middle German and Upper German areas from Ripuarian, Moselle-Franconian and Thuringian to Alemannic and Bavarian.

EXAMPLE OF AN ENTRY

The project started in 1994. From this date the main task has been the compilation of the corpus and database, as a result the final structure of each entry has not yet been clearly defined although general principles have been developed: the headword will be given in normalised spelling according to the *Mittelhochdeutsches Handwörterbuch* of Matthias Lexer (1872-1878); it will be followed by information about the word class/grammatical category, and a reference to the two existing Middle High German dictionaries, the one by Lexer and the *Mittelhochdeutsches Wörterbuch* by G.F. Benecke/W. Müller/F. Zarncke (1854-61). Some entries will be accompanied by information about orthographic variation, phonology and morphology based on the occurrences given in the entry. The main part of the entry will consist of a semantic and/or syntactic analysis of the usage of the headword, illustrated with examples from the corpus.

History of the Project and Planning

SHORT HISTORY OF THE PROJECT

The deplorable situation of Middle High German lexicography is generally well known. The *Mittelhochdeutsches Handwörterbuch* appeared in 1878, since which date all plans for a new dictionary covering the period have been unsuccessful. The main Middle High German editions, especially the series *Deutsche Texte des Mittelalters*, provided glossaries as a means of presenting to the scholar such material as was not included in the earlier dictionaries.

It was not until 1985 that a group of Trier researchers (K. Gärtner, C. Gerhardt, J. Jaehrling, R. Plate, W. Röll, E. Timm; Data processing: G. Hanrieder) started compiling the *Findebuch zum mittelhochdeutschen Wortschatz*, which was published in 1992. Of course, a number of dictionaries appeared which detailed the vocabulary and usage of individual authors or works, for example the *Wörterbuch zur Göttinger Frauenlob-Ausgabe* by K. Stackmann (1990) and the *Wörterbuch zu Veldekes Eneit* by G. Schieb (1970).

In the late eighties, on the basis of the work on the *Findebuch*, the present editors-in-chief began planning a new comprehensive, four-volume dictionary of Middle High German, in collaboration with the Academies of Sciences in Göttingen and Mainz. With the support of these Academies the current project was able to secure funding from the German Research Association (*Deutsche Forschungsgemeinschaft*) for the initial research period of six years. After that the Academies in Göttingen and Mainz are to finance and house the project.

(PLANNED) YEAR OF PUBLICATION

The project is a long-term one and the final published version will not appear for at least another 20 years.

FORM OF PUBLICATION

It is intended to make the corpus and database (*Belegarchiv*) available in a provisional form to scholars via the Internet. The final version will appear in print and most probably in electronic media (CD-ROM, Internet, etc.).

Lexicographical Tools and Methods

The two research centres in Göttingen and Trier are co-operating closely using the Internet for the everyday work and to discuss and exchange data and ideas.

HARDWARE

The hardware in use comprises a number of powerful PCs (DOS, Windows95, LINUX) and workstations (UNIX). The computing

centres of the Universities of Göttingen and Trier have provided the technology to secure the data on central servers.

SOFTWARE

The main software used for the project is TUSTEP (*Tübinger System von Textverarbeitungsprogrammen*). This system was specially devised as a tool in linguistic and literary data processing. It consists of several separate modules which allow a greater flexibility in dealing with particularly large sets of data than many other data processing systems. The processes devised for the project by P. Sappler (University of Tübingen) work with two main components: firstly, a thesaurus of lemmata and word-forms (*Gerüst*) gradually enriched with the addition of each new source; secondly, the text corpus, each individual text of which is prepared and pre-edited for processing using the thesaurus. This thesaurus is a key in lemmatising the prepared texts, producing, as a result, a concordance to each text. These individual concordances are checked for errors before the texts can be added to the body of prepared texts in order to create a general concordance, which is the database for the dictionary (*Belegarchiv*).

Special Remarks

NAMES

Anthroponyms and toponyms are not included. Elsewhere there are plans to produce a dictionary of proper nouns in Middle High German literature.

Ordbog over det Norrøne Prosasprog

A Dictionary of Old Norse Prose

Information about the Dictionary

EDITORS

Helle Degnbol
Bent Chr. Jacobsen
Eva Rode
Christopher Sanders
Þorbjörg Helgadóttir

ADDRESS

Njalsgade 76
2300 Copenhagen S
Denmark

tel.: +45 (35) 32-8490
fax: +45 (35) 32-8488

Contents of the Dictionary

TYPE

Ordbog over det norrøne prosasprog (*ONP*) is a scholarly, alphabetically-arranged, historical-semantic dictionary. Planned as a 12-volume work (the first volume and a volume of indices are already published), the c. 75,000 standard entries give information about grammatical status, inflection, oldest recorded occurrence and meaning. The latter is illustrated with quotations from manuscripts (unnormalised) and printed editions (normalised and unnormalised). The spelling of the headwords is normalised. Meaning is principally rendered by translation equivalents in both of the dictionary's metalanguages, Danish and English. The date of the manuscript copy of each work from which a citation is taken can be traced in the index volume via the abbreviation or siglum at the end of each citation.

SUBJECT/CORPUS

The dictionary covers the prose corpus of Old Icelandic and Old Norwegian. Runic inscriptions and poetry are not treated. In principle every vellum manuscript of each work and the paper manuscripts which text editors have assessed as possessing independent value as representatives of lost medieval exemplars are covered. Citations of this latter type from post-medieval copies are marked with a special indicator. In terms of genre the corpus includes every type of Old Norse saga: íslendingasögur (sagas of Icelanders), konungasögur (kings' sagas), fornaldarsögur (mythic-heroic sagas), byskupa- og samtíðarsögur (bishops' sagas and contemporary sagas), helgi- og postulasögur (hagiographic works and lives of the apostles) – mainly translated from Latin, see SPECIAL REMARKS below), riddarasögur (translated and original romances); in addition there are law texts and charters, encyclopaedic works, theological treatises, and annals. Some 750,000 tokens represented by the same number of dictionary slips will be treated. Place-names and personal names are not included. By-names are registered but only rarely interpreted.

DESCRIBED PERIOD

Old Icelandic is treated from the time of the oldest vellum manuscripts (c. 1150) to 1540 (the date of publication of the first Icelandic printed book) and Old Norwegian is covered from the time of the oldest vellum manuscripts to c. 1370.

DESCRIBED AREA

The Danish term 'norrøn' and the English equivalent 'Old Norse' (synonymous with 'West Norse') which are used in the title of the dictionary cover the language of the Faeroe Islands as well as of Iceland and Norway in the medieval period. There survive, however, precious few medieval documents that can be identified as Faeroese.

EXAMPLE OF AN ENTRY

(a masculine noun: á•verki)

Each standard entry in the dictionary may contain the following elements: the headword in normalised spelling and cardinal form (nominative singular for nouns, etc.); word class and gender; inflection; possible references to headwords that are morphologically closely-related; (for prepositions and verbs, some nouns and adjectives) an indication of case governance and typical syntactic combination patterns; an indication of meaning in both Danish and English; possible highlighting of collocations (printed in bold type) with a separate semantic analysis; a series of supportive quotations with abbreviated references (sigla); a list of recorded compound words in which the headword in question forms the last element; a list of earlier dictionaries and word-lists that have analysed or recorded the headword in question; a list of secondary literature which is relevant for the semantic analysis of the headword.

The oldest recorded citation is always given and it is individually marked. A symbol at the head of an entry indicates the cases where all available citations are quoted in the entry, and there is occasional recourse to "bare references" (sigla without quotations).

Alongside standard entries there are also secondary entries. These typically serve as a registration of a word's existence within

the defined corpus but omit the semantic analysis and supporting quotations, e.g. in the case of words recorded in poetry only or foreign words (as opposed to assimilated loanwords which are treated in standard entries).

History of the Project and Planning

SHORT HISTORY OF THE PROJECT

In 1939 the dictionary project was incorporated in the Danish state budget in the form of an annual subvention to Den arnamagnæanske kommission, "the Arnamagnæan Commission", which is responsible for the administration and publication of the completed work. The Arnamagnæan Commission is named after the Icelandic scholar Árni Magnússon (1663-1730), professor at the University of Copenhagen, who built up a large collection of Icelandic manuscripts, which he bequeathed to the university together with his books and the rest of his estate. During approximately the first 40 years, the project's limited staff was principally engaged in an excerpting of the corpus with a view to replacing the two major previous dictionaries from the nineteenth century, Johan Fritzner, *Ordbog over Det gamle norske Sprog* (Kristiania, 1886-1896) and Richard Cleasby & Gudbrand Vigfusson, *An Icelandic-English Dictionary* (Oxford, 1874, 2. rev. ed. 1957). The excerpting was selective. Each text was read by two editors who marked the items for which slips were subsequently written out by assistants.

In 1983 a prospectus was published and this was followed in 1989 by the volume of indices, which with its survey of the dictionary's corpus also forms an independent handbook to the study of Old Norse: *Ordbog over det norrøne prosasprog: Registre* (A Dictionary of Old Norse Prose: Indices) (Copenhagen, 1989). The first volume of the dictionary proper appeared in 1995: *Ordbog over det norrøne prosasprog 1: a - bam* (A Dictionary of Old Norse Prose 1: a - bam) (Copenhagen, 1995).

PUBLICATION PLAN

The remaining ten volumes are planned to appear at regular intervals over the next 28 years. There are no specific plans for

electronic publication, but it is a desideratum that the growing electronic database be made available on-line at some point in the future.

Lexicographic Tools and Methods

THE ARCHIVE

The dictionary's principal archive is a collection of some 750,000 handwritten dictionary slips (headwords and context) which are the fruits of the long period of excerpting. There is also an archive of compound words alphabetised according to the final element in each compound.

SOFTWARE: THE PRODUCTION AND PROCESSING OF DICTIONARY ARTICLES

The operational tool of the dictionary is an in-house designed relational database developed in the environment of a DOS shelfware package (SMART). At the start of a production run student assistants key an alphabetical series of slips into the database before checks, and, where necessary, updating of sigla (abbreviations indicating the source of the citations) are undertaken electronically. The resulting files are passed on to the editors who carry out all their activities in a series of inter-related database procedures. Finally a text file is extracted from the database and converted into $^{T}E^{X}$ (Postscript) format which generates camera-ready copy.

HARDWARE

As of yet, the hardware consists of DOS computers in a Novell local area network.

Special Remarks

SOURCE MATERIALS IN FOREIGN LANGUAGES

Foreign sources and parallels have been established for a considerable number of the texts that were translated into Old

Norse. The largest group is the hagiographic and theological material translated from Latin, but there is also a sizeable corpus of romances translated from Old French (Anglo-Norman) alongside legendaries translated from Low German, exempla derived from Middle English, and medical literature translated from Danish. Where relevant the readings of foreign sources or analogues are incorporated in the Old Norse citations in the dictionary.

[1]á·verki *sb. m.* [-a; -ar] *(cf.* áverk *sb. n.,* áverki *sb. n.,* áverkr *sb. m.)*

1) *legemsskade, legemsbeskadigelse // bodily injury, wound:* hjó hann til Steins stýrimanns ok veitti honum mikinn áverka *Gr 30[24];* oc do sialfr af uheyrðum harmkvælum, vall moðkum oc veitti ser averka með knifi *JBapt[2]A 866[38];* lavst hann með avxar-hamri, oc vard þat synn averki *Stu[1]K 71[6];* hjó Kjartan fót af Guðlaugi fyrir ofan kné, ok var honum sá áverki œrinn til bana *Laxd 190[18];* Ec lysi sár þav mer a hond. oc þaN a verka allan er a honom er uNiN *GrgKon[1] 154[2];* þeir luki alt sem dœmt er loglega h<v>arium ser i lage sem auerkan laut i hvari tigvnd *RbHM (*1307)[2a] 72[29]; item: GrgKon[II] 73[27]; StjœrnODr[x] 107[1a].*

2) [á e-m / við e-n] *overfald, angreb // assault:* Hefe ec synþer gort i morþe oc i maNdrape. J averkom a cristnom mœNom • *Játn[1] (1993) 67v[10];* veittu þeim manni averka er Jngialldr het ok skutu hann *HákFlat 127[32];* hann skyldi vita, ef hann mætti leynask inn í skálann ok leita þaðan til áverka við menn *Eb[x] 116[8]; item: Stu[1]K 541[27]; TrójHb 30[26]* **sæta áverkum** [við e-n] *overfalde, angribe // assault, attack:* Snero þeir Kalfr þegar at moti þeim oc sættv averkvm við Hall, enn hann horvaði vndan oc varði sig *Stu[1]K 295[23];* ætla ek þér, Már frændi, at sæta áverkum við Bjorn *Eb[x] 133[14]; item: DplFlj[x] 125[4]; GullÞ 37[2]* **áverkar verða** [með e-jum] *slagsmål opstår (mellem nogle) // a fight starts (between people):* verþa nv j borgiNNe stvNNdvm averkar meþ moNvm oc æinstaka vig *Jvs291 85[15];* urðu áverkar með mønnum um daginn *Eb[x] 107[17];* komst Aron jnn j skemmuna. og .iiij. menn med honum: ok uerða auerckar ok þa fieck Tumi sar mikit *GBpA394 195[28]; item: HákFlat 33[28] (cap.); HrHalt 412[14]*

3) [á e-m] *erstatning for legemsbeskadigelse // compensation for bodily injury:* Huervetna þar sem maðr a a ser averka at hæimta. eða a þæim umaga eða æigi fultiða manne sem maðr a umboð a ... þa *Bl 211[26]*

4) *udbytte/afkast (af jordbrug), ?byggeri (på gård) // produce/profit (of farming), ?building work (on a farm):* Nv skal gorðum skipta eptir iarðar hofn oc allum auerka *Landsl 122[2];* ek lauk Þorere Erlendzsyni halft fiordæ laups boll jerder ... meder lunnyndum en till hafuæ leghet vttan gardz ok jnann ok meder ollum auærkæ sem mit hafuum a giort sidæn en daughen (!) var *DN III (1358) 240[19];* Abuð iarðar heimilar tekiu en landskylld heimilar loð oc allan auerka þann er i kaup þeirra kom *Landsl 106[1]*

5) [á e-u] *beskadigelse ved uretmæssig udnyttelse af anden persons jord // damage in conn. with illegal use of another person's land:* þau áren sem minne averki hefer ordit áá jordunne luki aufundarbot epter domi *DI V (1474) 741[17] (cf.* sem hann hefer ... skada giort áá kirkiunnar jordu *741[14-15]);* Allt þat er menn skil a vm auerka a akre eða eng *Landsl 119[5];* En ef hann stendr hann eigi *(at verki).* þá leiði menn til stoms *(ɔ: stofns)* oc sýni áverca oc láti meta viðarspell oc kenni þeim er hann vill *Frost[x] 243[25]; item: IslDipl (1401) 143[10]*

6) *udbytte/produkt ved uretmæssig udnyttelse af anden persons jord // product of illegal use of another person's land:* Hvervitna sem maðr hittir averca i morco sinni. þa scal hann taca brott at usecku *Gul[I] 45[6];* Nu a sa averca er iorð hever leigðt *Gul[I] 39[6]*

Comp.: alþingis-; skógar-

Gloss.: *ClV; Finsen* (áverk)*; Fr; Hertzb; Rím; AJ; Fr4; NO; ÁBIM; (Bl)*

Litt.: *Taranger 1896 108; Lúdvík Ingvarsson 1970 273-276*

Gammalnorsk Ordboksverk c. 1200-1550

Dictionary of Old Norwegian c. 1200-1550

Information about the Dictionary

EDITORS

Magnus Rindal (on leave 1993-1999, in charge of the University of Oslo Centre for Viking and Medieval Studies)
Erik Simensen
Bjørn Eithun

INSTITUTE

Institutt for nordistikk og litteraturvitenskap

ADDRESS

Universitetet i Oslo
Institutt for nordistikk og litteraturvitenskap
Avdeling for leksikografi
Gammalnorsk ordboksverk
Postboks 1001 Blindern
N 0315 Oslo
Norway

e-mail: Erik.Simensen@inl.uio.no
Bjorn.Eithun@inl.uio.no
tel.: Erik Simensen +47 (22) 85-4386
Bjørn Eithun +47 (22) 85-4323
fax.: +47 (22) 85-4381

Contents of the Dictionary

TYPE

Gammalnorsk ordboksverk is a linguistic (word) archive containing at present about 600,000 individual slips registering occurrences of words in Norwegian medieval documents. The sources have been excerpted according to the thesaurus principle, i.e., every individual occurrence is registered separately. Each word slip contains a normalised lemma form, the particular manuscript spelling, information about the word class and a determination of the grammatical form, a specification of the manuscript source, and a short section of the text (typically about five lines) showing the word in context. Nouns and pronouns are listed as lemmas in the nominative singular form, verbs in the infinitive. Lemmas are sorted alphabetically, and are grammatically ordered within each word class. Occurrences of words in official documents (charters) or diplomas are also ordered according to date.

SUBJECT MATTER

The corpus of source material consists of all documents in vernacular Norwegian from the first extant record written with the Latin alphabet in Norway in the early High Middle Ages and up to the time of the Reformation. Runic inscriptions are not included. The considerable number of diplomas concerning Norwegian affairs that are written in Latin and Low German are not included in the archive. As work progresses, special procedures will have to be devised for treating the increasing number of documents in the other Scandinavian languages – Swedish and particularly Danish – towards the end of the period. The corpus may be divided into three main categories: 1) literary texts (saga texts in a wider sense); 2) diplomas, and 3) legal texts. Most of the relevant literary texts are included in the archive itself. The *Gammel norsk homiliebok* (Old Norwegian Book of Homilies, manuscript AM 619 4to) and the earliest fragments of other texts are, however, passed over since they were excerpted and printed as a thesaurus in *Ordforrådet i de eldste norske håndskrifter til ca. 1250,* ed. Anne Holtsmark (Oslo, 1955). Diplomas and in particular laws have been treated to a much lesser degree lexicographically. This situation is being rectified as a

result of the present computerisation of Old and Middle Norwegian source material under the auspices of the so-called Documentation Project. Most of this material has been published, but only in non-diplomatic or not strictly diplomatic editions, like those in the *Diplomatarium Norvegicum (DN)*, vols I-XXII (with approximately 1000 pages per volume), and in *Norges gamle Love* (The Ancient Laws of Norway), series I, vols I-V. In spite of its semi-diplomatic form, the *Diplomatarium Norvegicum* is being digitised. Of the law texts, only a minor, but important part – *Gulatingslova* (The Old West Norwegian Law Code of Gulating) – has been digitised; the text files have, however, not yet been loaded into a database. The same applies to *Magnus Lagabøtes Landslov* (in the Eidsivating redaction; The National Law Code of King Magnus the Law-Mender, manuscript AM 309 fol.). The situation is similar for *Baarlaams ok Josaphats saga* (The Saga of Barlaam and Josaphat, manuscript Holm. perg. 6 fol.).

In the course of 1997, according to present plans, most of the archive will be electronically accessible.

DESCRIBED PERIOD

The Old Norwegian language period has traditionally been divided into two parts, before and after the year 1350 (the year of the great pestilence; or 1370, when the consequences of the plague became clearer). The earlier part is usually called the period of (classical) Old Norwegian, whereas the latter part is termed the Middle Norwegian period. Until recent years, the work of excerpting has been concentrated on the Old Norwegian period. Be that as it may, the objective of *Gammalnorsk ordboksverk* is to build and maintain a complete word archive – a thesaurus – containing the vocabulary from all Norwegian documents written with the Latin alphabet from the very first item (c. 1200) up to the time of the Reformation (about 1550).

DESCRIBED AREA

Norwegian sovereignty during the Middle Ages at times encompassed vast areas in western waters from the British Isles to Greenland. However, only documents from the Orkneys, Shetland and the Faeroe Islands, and some areas in western districts of

present-day Sweden (i.e. Jämtland, Härjedalen, Idre, Särna and the coastal district of Bohuslän) have been included. Documents in Old Icelandic are not included in the corpus.

History of the Project and Planning

SHORT HISTORY OF THE PROJECT

Preparations for *Gammalnorsk ordboksverk* began in 1934, and discussions took place with colleagues in linguistic fields in Copenhagen, who were planning an Old Norse (i.e. Old Norwegian and Old Icelandic) dictionary. In 1937 the Danish initiative led to the establishment of "Den Arnamagnæanske Kommissions Ordbog", now being published under the title *Ordbog over det Norrøne Prosasprog* (A Dictionary of Old Norse Prose). In 1940, *Gammalnorsk ordboksverk* started quite modestly, with the first grants from private funds. Further co-operation with the project in Denmark was made impossible by the war, and work continued independently in Copenhagen and Oslo. A joint meeting in 1947 concluded that two different projects had developed, and that it would be best if they continued to be run along their respective chosen lines. Several Norwegian scholars contributed to the formation of the Norwegian project, in particular Ludvig Holm-Olsen, librarian at Oslo University Library, later professor at the University of Bergen, Anne Holtsmark and Didrik Arup Seip, professors at the University of Oslo. The lexicographical task was a part-time enterprise undertaken above and beyond ordinary professional duties. The tasks of excerpting texts, producing word slips, writing lemmas, and other necessary routine work were performed for the most part by hired part-time senior students and other qualified personnel. Finn Hødnebø was the first lexicographer engaged full-time at the *Gammalnorsk ordboksverk* (from 1950). In 1972 *Gammalnorsk ordboksverk* and two other Norwegian dictionary projects established "Norsk leksikografisk institutt" (Department of Norwegian Lexicography), which was the first and only lexicographic institution in Norway with university affiliation. The formal status was again changed in 1990 when the institute, as a result of a major university reorganisation, lost its status as an

independent university department and became the "Avdeling for leksikografi" (Section for Lexicography) under the Department of Scandinavian Studies and Comparative Literature. *Gammalnorsk ordboksverk*, however, retains its traditional name.

PRESENT SITUATION

Technical approaches and daily routines in converting linguistic evidence into purposeful formats have constantly been changing. In the 1980s the introduction of computers brought new challenges and put an end to many old routines. Traditional word slips will most likely not be produced any more.

Technology has also created new problems, which have to be addressed. A more professional attitude was adopted when in 1991 *Gammalnorsk ordboksverk* became part of the University of Oslo Documentation Project, "which aims at applying modern computer technology to the collections of language and culture". The project was later extended to include all four universities in Norway and renamed "the University Databases for Language and Culture". A primary aim of the lexicographical part of the Documentation Project is to co-ordinate the various word archives and collections in a comprehensive information bank. According to plans, this resource-demanding project in its present form is to be completed by the end of 1997.

About half of the paper-slip archive in the *Gammalnorsk ordboksverk* had been computerised by the end of March 1997, and in addition most of the twenty-one printed volumes of the *Diplomatarium Norvegicum (DN)*. However, a considerable amount of control work and final proof-reading still remains to be done. A WorldWideWeb searchable test database of the *DN* has been made accessible at the following address (for the time being with explanatory text in Norwegian only): http://www.dokpro.uio.no/cgi-bin/cesoS/diplom_felt

The earliest of the documents in this database date from the 990s (some in Old English), the latest from the 1580s (mostly in Danish). All diplomas published in *DN* are included, also those in Latin and Low German.

COMMENTS ABOUT THE FUTURE

We are confident that the future word database will improve the accessibility of information about the Norwegian language, history and culture for the benefit of research, teaching and general knowledge.

The experience and results acquired through the Documentation Project will be useful for the future work routines in *Gammalnorsk ordboksverk.* Digitised versions of elder non-diplomatically printed editions of diplomas and law texts such as those in the *Diplomatarium Norvegicum* and *Norges gamle Love* can most likely serve as a basis for conversion of such material into a diplomatic version. Routines for at least partly automatic grammatical marking of unmarked texts will be generated and developed, based on the new SGML-tagged and digitised word archive.

The use of electronic databases will in the future be wider and more versatile than has been the case with the traditional archive. Whereas there are at present no plans for producing a printed dictionary of Old Norwegian or other dictionary of Old Norse based upon this database, initial preparations have been made for the making of a sorely needed dictionary of Middle Norwegian (see "The Dictionary of the Written Language in Norway 1350-1550" elsewhere in the present volume).

Some trials of digitisation of images have been carried out. Test scannings of photographs kept in *Gammalnorsk ordboksverk* of medieval diplomas and manuscripts have given interesting results. Links between text databases and image databases open new perspectives and point to new procedures for the handling of medieval texts. Please refer to the test figure at the end of this article.

Lexicographical Tools and Methods

THE HARDWARE

The new era of digitisation did not start very professionally, but rather more at random, and the initiative was taken by the editors

themselves with little knowledge in this field and little external assistance. One basic idea in the beginning was simply to substitute computers for typewriters and stencil machines in the production of word slips for the traditional archive. This procedure would give the additional bonus of creating a digitised corpus with powerful inherent capabilities.

Computerisation is now entirely different from what it was in the early 1980s, and the Digital DEC and VAX mainframes have given way to the present-day powerful personal computers. For a short intermediate period an eight-terminal microcomputer, US brand North Star, model Horizon, constituted our hardware, but this machine was later replaced by two IBM PC AT computers. From the start of the Documentation Project in 1991 some IBM PC models and a variety of Apple Macintosh computers have been employed.

THE SOFTWARE

Computerisation started as a simple and straightforward approach to the process of producing word slips. Text was basically entered in a fashion similar to using a typewriter, but now the CP/M operative system and the word-processor WordStar were employed. Old Norwegian special characters had to be handled in a new way, and a home-made coding system was developed (which produced, however, cryptic printouts and made proof-reading laborious). With the IBM machines, the new operative system (starting with DOS 3.1), and new word-processors (starting with WordPerfect 4.2), the computerisation was continued. With the Macintoshes, yet another system and other word processors, Claris MacWrite II, and MS Word (starting with version 4.0) were introduced. Some registration work is also being done by means of the Apple application HyperCard and FileMaker Pro from Claris Corporation. Different file formats are now being converted to the Unix format. More professional database applications will be developed by the programming group of the Documentation Project.

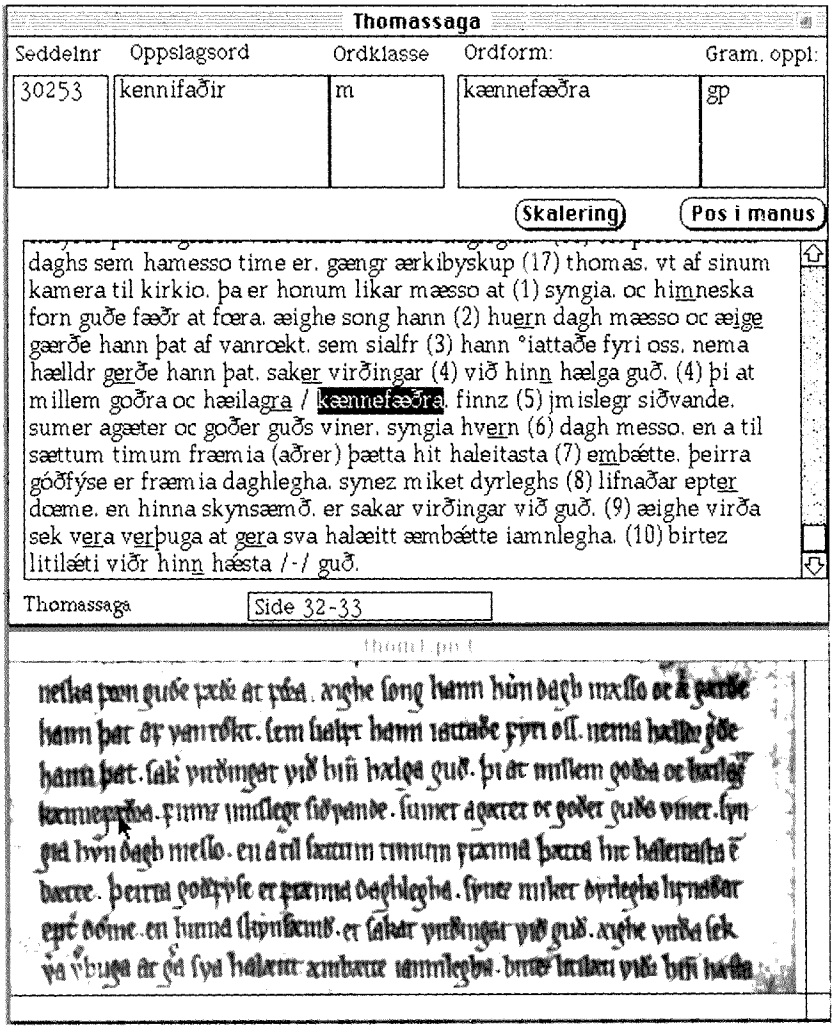
MANUSCRIPT IMAGES

Gammalnorsk ordboksverk continues to build a collection of photographic documentation of medieval manuscripts. The images

are mainly full-size black-and-white photographic enlargements processed from 9x12 cm sheet film or 35 mm black-and-white roll film. Colour diapositives constitute a small part of the collection, and a few documents have been converted to the Kodak PhotoCD format.

There is some uncertainty as to which approach should be taken toward digitisation of manuscript images. Present scanning technology encompasses the drum scanner and the flatbed scanner. Unfortunately, the former, which is the best scanning alternative, is hardly viable for old and fragile documents; the technology of flatbed scanners is, however, constantly improving. Even better results can be achieved by employing high-class digital cameras like those manufactured by Kontron Elektronik in Germany.

With the present state of the art, the best approach to digitisation of medieval manuscripts would be the re-photographing of the originals with first-class digital cameras. However, the existing collections of medieval document images are so large that much can be said for converting the existing photographic images to digital form. Picture quality is in general quite good, and these photographs constitute a good basis for a temporary solution. Top-end flatbed scanners could manage the digitisation of these images at a relatively low cost and high speed. Such a procedure should, however, not preclude the pursuit of the preferred approach. A new round of high-class medieval manuscript photography would be a very time-consuming project, and would call for systematic planning and good co-operation between the partners involved.


Thomassaga
Seddelnr
Oppslagsord
Ordklasse
Ordform:
Gram. oppl:
30253
kennifaðir
m
kænnefæðra
gp
Skalering
Pos i manus
daghs sem hamesso time er. gængr ærkibyskup (17) thomas. vt af sinum kamera til kirkio. þa er honum likar mæsso at (1) syngia. oc himneska forn guðe fæðr at fœra. æighe song hann (2) huern dagh mæsso oc æige gærðe hann þat af vanrœkt. sem sialfr (3) hann °iattaðe fyri oss. nema hælldr gerðe hann þat. saker virðingar (4) við hinn hælga guð. (4) þi at millem goðra oc hæilagra / kænnefæðra. finnz (5) jmislegr siðvande. sumer agæter oc goðer guðs viner. syngia hvern (6) dagh messo. en a til sættum timum fræmia (aðrer) þætta hit haleitasta (7) embætte. þeirra góðfýse er fræmia daghlegha. synez miket dyrleghs (8) lifnaðar epter dœme. en hinna skynsæmð. er sakar virðingar við guð. (9) æighe virða sek vera verþuga at gera sva halæitt æmbætte iamnlegha. (10) birtez litilæti viðr hinn hæsta /-/ guð.
Thomassaga
Side 32-33

Example of linked text database and image database. From the *Saga of Archbishop Thomas* (Parchment quarto no. 17 in the Royal Library in Stockholm).

Short translation of terms used:

Seddelnr. - Word slip number
Oppslagsord - Lemma
Ordklasse - Word class
Ordform - Manuscript word form
Gram.oppl. - Grammatical information
m - Masculine gender
gp - Genitive plural
Skalering - Image scaling
Pos i manus - Word position in the manuscript
Side - Page (in the printed reference)

Ordbok over Skriftspråket i Norge 1350-1550

Dictionary of the Written Language in Norway 1350-1550

Information about the Dictionary

EDITORS
Erik Simensen
Bjørn Eithun

INSTITUTE
Institutt for nordistikk og litteraturvitenskap

ADDRESS
Erik Simensen
Universitetet i Oslo
Institutt for nordistikk og litteraturvitenskap
Avdeling for leksikografi
Gammalnorsk ordboksverk
Postboks 1001 Blindern
N 0315 Oslo
Norway

e-mail: Erik.Simensen@inl.uio.no
tel.: +47 (22) 85-4386

Contents of the Dictionary

TYPE

The *Dictionary of the Written Language in Norway 1350-1550* (*DWLN*) is an alphabetically ordered, diachronic period-dictionary. The editorial language is Modern Standard Norwegian (*Bokmål*). The entries (see below) give information on morphological status (word class, inflection, affixes and/or compound elements when relevant), number of occurrences, meaning, oldest attestation (when the item in question is not known from Old Norse prior to 1350), and etymology. The spelling of the entries is normalised. Nouns and pronouns are given in the nominative singular, verbs in the infinitive.

SUBJECT/CORPUS

The dictionary is based on a corpus of 1) diplomas (charters); 2) law texts, and 3) land registers, provided that the texts are written (with the Latin alphabet) in a Nordic language in what was then Norway (see below). The corpus will be selected from a total of about 19,000 printed pages, most of which have been edited in *Diplomatarium Norvegicum* (up until now 22 volumes), *Norges gamle Love* (The old laws of Norway), and the land registers associated with some prominent bishops (Aslak Bolt, Øystein Aslaksson and Olav Engelbrektsson).

DESCRIBED PERIOD

The period the dictionary describes begins with the year 1350, which, being the year after the Black Death came to Norway, has become almost symbolic as a mark of transition to a late medieval form of Norwegian. (This reflects a commonly – but not by all scholars – accepted division of Norwegian language history.) It ends with the year 1550, at which time the transition into Modern Norwegian is supposed to have been finished (at least in principle, and especially on the phonological level).

DESCRIBED AREA

The area comprises what was then (1350-1550) considered to be the realm of Norway, i.e. the present Norway and the (now Swedish)

regions Bohuslän (in south-western Sweden), Jämtland and Härjedalen (bordering on Mid-Norway).

STRUCTURE OF THE ENTRIES

In the printed version, each entry in the *DWLN* contains the following information: first the headword, in normalised spelling, followed by the word class; next come the total number of occurrences of the headword in the corpus, then important spelling variants, attested inflection, remarks on etymology, and subsequently meaning(s): thereafter phraseology, and finally information about the geographical distribution of the headword. In addition, the corresponding entry in the computer base will list the textforms of each headword, the immediate linguistic context, and the sources of all the textforms.

History of the Project and Planning

SHORT HISTORY OF THE PROJECT

The initiative was taken by professor Magnus Rindal (University of Oslo) with a memorandum in 1987, and some preliminary excerpting was done by E. Simensen in 1991. Since then the project has been in abeyance for some years, awaiting the progress of the Documentation Project (1991-1997, hereafter: DP), a large-scale undertaking (see the article of Bjørn Eithun, elsewhere in this book) consisting in the full computerisation of the Norwegian dictionary archives and great source publications (see above) relevant to the making of the planned dictionary. When this goal is achieved, the corpus can be selected from the database and tagged with the desired information. Due to considerable uncertainty with respect to the resources and progress of the DP in its last phase, no fixed year of publication has been scheduled.

FORM OF PUBLICATION

The *DWLN* will be published in print and also be available on CD-ROM. The printed version is planned to be a one-volume dictionary of about 800-1000 pages, with a limited amount of reference to its sources. The introduction will give a short history of the project, an

outline of the general characteristics of the relationship between graphemes and phonemes, along with an account of principles for normalisation and editorial practice.

Lexicographical Tools and Methods

THE SOFTWARE

In the preliminary stage the simple software of FileMaker II (copyright Claris Corporation, Mountain View, Calif., USA 1988) has been used, mainly for practical uses. At present other, more powerful database programs are being considered. A possible production procedure is sketched in the Appendix below.

Special Remarks

NAMES

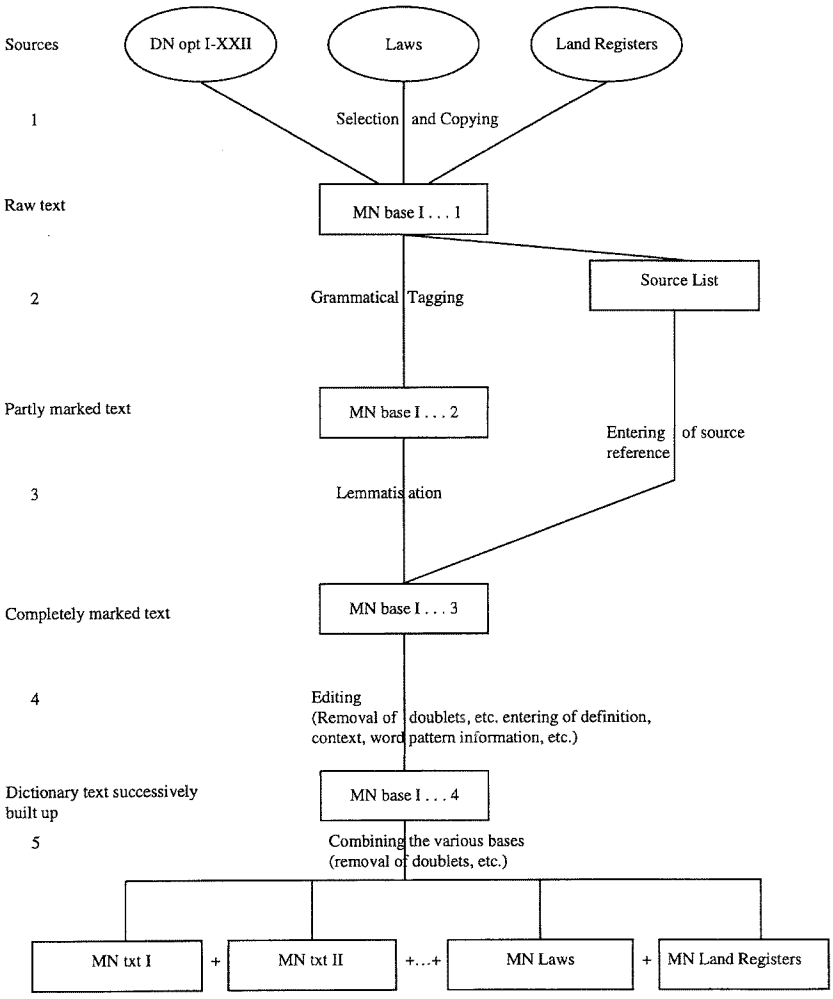
As a rule, anthroponyms and toponyms will not be included, since they are already collected and published elsewhere, by E.H. Lind (personal names) and O. Rygh et al. (place names, especially farm names). Exceptions may be made for latinised forms of Nordic names (e.g., Norvegia) and Nordic forms of foreign names (e.g. Lundunir, Øxnafurda).

LATIN/NORWEGIAN

Foreign words, especially of Latin origin, will only be included if they are inflected according to the grammar of a Nordic language.

Appendix

DWLN 1350-1550
Sketch of a Possible Production Procedure



II Papers

On the Preparation of a Dictionary of Middle Norwegian

Erik Simensen

This article addresses two sets of questions. Firstly, the concept "Middle Norwegian" – what it is, or rather what we think it is, and whether it can be justified as a useful term in the description of Norwegian language history; secondly, why we need a dictionary for the language period traditionally called by that name.

Since Norwegian is one of the minor European languages, it may be appropriate, by way of introduction, to explain briefly the concept of "Middle Norwegian".

1. The term "Middle Norwegian" refers to the Norwegian language of the late Middle Ages, roughly from the time of the Black Death (1349) or a little later (c. 1370) until the Lutheran Reformation (c. 1525). It is not sharply distinguished from the classical Old Norwegian of the thirteenth century. It is rather a period of gradual transition from Old to New (Modern) Norwegian, reflecting a general development from a synthetic to a (weakly) analytic type of language. Some of the important changes taking place in this period are the following:

(1) [þ] is replaced by /d/ or /t/
(2) The distinction nominative vs. accusative disappears in nouns
(3) Adjectives lose their case inflection
(4) Personal endings vanish from verb inflection

(5) More loan-words and calques come into the vocabulary, mainly from Danish, Swedish, and Low German[1]

This period of transition starts already in the thirteenth century, and some of the changes involved, e.g., the Quantity Shift (in the stressed syllables), are still not completed in all Norwegian dialects. This is the reason why a distinguished Norwegian scholar, professor Magnus Rindal, feels that the grounds are insufficient for honouring these two centuries with the name of 'period', with an appellation of their own. Instead, he stresses the continuity of Old Norwegian throughout the Middle Ages, but finds it reasonable, in view of the many innovations in the thirteenth century, to distinguish (1993:401) between the classical Old Norwegian (c. 700-c. 1250) on one hand and on the other hand late or younger Old Norwegian (c. 1250-c. 1500). Nevertheless, the fifteenth century is characterised by more intense competition between older and younger linguistic forms.

2. The second point mentioned above concerns the need for a dictionary, preliminarily, and for short, called a "Dictionary of Middle Norwegian". It is a sad fact that the later stages of Old Norwegian are poorly covered lexicographically, especially the late fifteenth and the sixteenth century. Fritzner's dictionary goes down to about 1450, Cleasby/Vigfusson somewhat further, but mainly for Icelandic sources. There exists a glossary to law texts down to 1447,[2] and some lists of loan-words in treatises by M. Hægstad and P. Grøtvedt. That is just about all there is.

The Norwegian language of these Dark Ages requires a description, for which a dictionary is an indispensable precondition.[3] Apart from historically oriented linguists, who else will benefit from the existence of such a dictionary? Mainly the historians, and others working with texts from the later Middle Ages. These texts are often difficult to read and understand. Firstly, they contain many new words

[1] Readers who want to know more about this period can be referred to Einar Haugen's book *The Scandinavian Languages*, chapter 11.

[2] Johnsen 1912.

[3] This need has been expressed earlier, for instance by Anne Holtsmark (1951:367): "Vi mangler f.eks. en mellomnorsk ordbok; mellomnorskens grammatikk er ikke skrevet [We still need a dictionary of Middle Norwegian; the grammar of Middle Norwegian has not been written]". Alexander Jóhannesson (1953:10) has made similar remarks with respect to the Icelandic language of the late Medieval Age.

and phrases, especially loan-words, not entered in existing dictionaries of Old Norse or Modern Norwegian. Secondly, they contain many words which were already known in Old Norse (perhaps 80-90 % of the vocabulary), but which are difficult to recognise and identify because they are written in a varying and inconsequent orthography. In some cases these words may be deciphered with the help of dictionaries of Older Danish (Kalkar) or Older Swedish (Söderwall), or Middle Low German (Lübben, Lasch) or New Norwegian (Aasen, Ross, Hellevik et al.). But some of these dictionaries are not easily accessible to non-Nordic scholars.

We need to catch those forms of Norwegian which have until now not been treated adequately from a linguistic and lexicographic point of view. A specimen of such a text is the following extract from a document from Østfold (in south-eastern Norway) 1526 (issued as DN III 1104):

> ... wi gøre idher kwnnocht ath mid worom i Hobøl ... sogom oc hørdom iaa oc handherbandhe *theres* aff eine halffwe her Mons Jon son, æn aff andre halffwe Gwttorm Swen son, ad the ware wm ens wm en gordh, som ... Gwttorm hadhe *solth* ... her Mons ... saa mykyn parth han otthe i eystre gordhen ... oc i *eydhegordhen* som heiter Breckan ... Mons skall haffwe dannemen ther offuer for^{de} *iordhe parthe* oc ransake bordtalet ... war soo i *theres wilkor* ad kyrken skall wpbere ... i for^{de} *eigedeildhe* ... xii oc xx norske mark *belateth* i for^{de} *iorde parthe*, som ... Gwttorm *loth* segh wel ad nøgie. Till meire wissan *trycke* wi wor incigle fore thette breff.[4]

Here are words that cannot be found in our dictionaries of the older language, e.g., *eydhegordhen* ('the deserted farm'), *eigedeildhe* ('property'), *iordhe parthe* ('parts of farms'). These words are probably results of autochtonous developments and are easily understood in the light of previous as well as later words, cf. Fritzner under *eyði-* and *eigu-* and modern dictionaries under *gård* and *jord*. Newer forms of well-known older words are *solth* ('sold'), *loth* ('let') and *theres* ('theirs') (instead of *selt*, *lét* and *þeira*). A few Middle Low German loan-words

[4] The italicised words in this and the following quotations are those which are commented on below.

have crept in: *wilkor* ('condition'), *belateth* ('left, given'), and *trycke* ('press').

Then the question arises: how do we delimit the scope of the material, and the time span to be covered? From the discussion above it will be remembered that the whole concept of Middle Norwegian has a somewhat dubious status as a label for a definite stage in our language history. For this and other reasons we have chosen partly non-linguistic criteria. Thus we take into consideration all documents written (with the Latin alphabet) in a Nordic language in what was then Norway (i.e., including Bohuslän, Jämtland and Härjedalen) in the period 1350-1550. That is why we have given this project the neutral title, *Dictionary of the Written Language in Norway 1350-1550.*

Behind the formulation "Written Language in Norway" lies the problem of what should legitimately count as Norwegian in this period. The problem is highlighted especially in texts with (presumably) foreign word forms, such as in a document from Hamar (Eastern Norway) 1466 (issued as DN III 873):

> ... kunnikt gørande thet *iak* war j hallinne i biskups gardhin i Hambre ... badha bygderna laang-samligha tretthe hafwer *warith* ... konigx vmbodsman *hulkin* tha ær næmpna the men som thet skal ransaka ... naar the *til hopa* koma ...

Some words here carry a definite Swedish ring, e.g. *iak* ('I'), *warith* ('been'), *hulkin* ('who') and *til hopa* ('together').

Another document, from Gudbrandsdalen (Mid-Norway) 1527 (issued as DN III 1106), seems influenced by Danish, e.g. in the words *fehaffue* ('enclosed pasture'), *bewisningh* ('proof, evidence'), *skutzmoll* ('reference') and *sandheid* ('truth'):

> ... som the vore lagligh tiith steffde oc framkraffde for nogen skogh oc *fehaffue* ... hulcked proffe breff eller *bewisningh* ... bøødz them *skutzmoll* ... At saa y *sandheid* er ...

These two documents have a predominantly Norwegian character (in phonology and morphology). But many texts leave one with the impression that they can hardly be said to be Norwegian at all. Indeed, part of the rationale of this project is to help us clarify what is really Norwegian in this period (1350-1550). Because we do not know for sure, we have to choose a fairly broad definition of "Norwegian", in order to avoid circularity and not to miss information.

3.1. The general character of the dictionary project is described elsewhere in this volume. Here, only a few additional remarks will be made. The sources are mainly three types of texts: 1) diplomas (charters); 2) law texts, and 3) land registers, e.g. those associated with some prominent bishops (Aslak Bolt, Øystein Aslaksson and Olav Engelbrektsson).

Lexically, the law texts are the easiest part of the task, because most of the vocabulary is known from classical Old Norwegian and already largely explained in the Glossary of *Norges gamle Love* (Norway's Old Laws [prior to 1447]). Our job will then be to select those which fall within our period and enter the more important deviant forms from younger transcripts of older texts. The land registers are voluminous, but their vocabulary is very standardised and will probably to a large extent be covered by the diplomas, which by far constitute the greatest part of the source material. For all three groups taken together it will be necessary to make selections within a corpus of c. 19,000 pages. As already indicated we have decided upon a certain chronological and geographic frame. We do not intend to publish a concordance. We plan to register all lexemes in the texts, all inflectional forms, and a selection of variants.

How much we should include of variant readings will be based on two considerations: 1) the presumed knowledge of the general users – we take it for granted that they have a basic knowledge of Old Norse, and 2) the contents of the introduction to the dictionary – it is supposed to outline the general characteristics of the relationship between graphemes and phonemes, along with an account of our principles for normalisation and editorial practice.

The rules for normalisation (of the headwords) cannot be determined until a bigger and more representative selection of the material has been studied. Our aim is to produce a diachronic, passive dictionary. "Passive" here means that phraseology and context, synonyms and equivalents will be restricted to a minimum. Our editorial language will be Standard Modern Norwegian (*Bokmål*).

A possible structure of the articles, as they will preliminarily be stored in the computer, is illustrated by the following examples of a representation format.

ENTRY	**øydigarder**
WORDCLASS	m.
TEXTFORM	eydhegordhen
TEXTGRAMMAR	best. akk. sg. [def. acc. sg.]
CONTEXT	
MEANING	ødegård ['deserted farm']
SOURCE	III 1104, Hobøl 1526
COMP. and/or DER.	+
OTHER INF.	

ENTRY	**trykkia**
WORDCLASS	v.
TEXTFORM	trycke
TEXTGRAMMAR	pres. pl. ind.
CONTEXT	
MEANING	trykke, presse ['press']
SOURCE	III 1104, Hobøl 1526
COMP. and/or DER.	–
OTHER INF.	

This format shows what kinds of information we primarily look for. Some categories (the first three + MEANING) apply to all words, others only to certain types. COMP[osition]. and/or DER[ivation]. is the place where compositions and derivations with certain specified affixes will be registered separately. As far as verbs are concerned, the example with *trykkia* is meant to show that verbs are here not inflected for person, although the number distinction (sg. vs. pl.) is kept. The verb forms *sogom* and *hørdom* (see the first excerpt above) appear to be exceptions, since the regular plural-ending in this particular text is *-e* (*tryckie*, *sendhe*, *gøre*, *ware*). The *-om*-forms are probably relict forms, taken over as imitations of a traditional formula in this type of text.

This invites the following reasonable question: why exactly are these kinds of information selected, and not (also) others? Our answer is that we think these categories will serve a purpose in providing materials for a grammar of Norwegian from this period. We badly miss such a grammar. This dictionary can make it possible to write one.

3.2. When comparing this dictionary project with two other more advanced projects, *DONP* (i.e. *Dictionary of Old Norse Prose*, Copenhagen) and *VMNW* (i.e. *Vroegmiddelnederlands Woordenboek*, Leiden), we find some striking differences. Let us look at the more important ones, and at the *DONP* first.

Firstly, the *DONP* is based on selective excerpting. This is understandable in the light of the time-span that the dictionary covers (c. 1100-c. 1540) and its vast scope of material. Secondly, quotations are equipped with source references, sometimes traced back to foreign originals (in the case of translations, etc.). Thirdly, the editorial text (i.e. translations and definitions) are given in two languages, Danish and English, which makes the dictionary easier to use for scholars who do not master a Nordic language.

On this point our dictionary will resemble the *VMNW*, in that it uses the national language of its country (Norwegian, like the *VMNW*'s Dutch) as its editorial language. The main differences are, as far as can be inferred from the fifth draft version of the *VMNW* (1993), that the *VMNW* also includes 1) names and 2) source references.[5]

[5] I wish to thank Mrs Oddrun Grønvik for her valuable assistance with questions concerning English usage.

References

Alexander Jóhannesson. 1953. Om det islandske Sprog. *Scripta Islandica* 4 : 5-14

The Arnamagnæan Commission (ed.). 1995. *Ordbog over det norrøne prosasprog. ONP 1 : Nøgle // Key*. Copenhagen

The Arnamagnæan Commission (ed.). 1995. *A Dictionary of Old Norse Prose. 1 : a-bam*. Copenhagen

Brinchmann, Christopher & Johan Agerholt (eds.). 1926. *Olav Engelbrektssøns jordebog*. Oslo

Cleasby, Richard et al. (eds.). 1957. *An Icelandic-English Dictionary*. 2nd ed. Oxford

DN: *Diplomatarium Norvegicum*. Christiania / Kristiania / Bergen / Oslo 1849. (Ref. to volume and number.)

Fritzner, Johan. 1886, 1891, 1896, 1972. *Ordbog over Det gamle norske Sprog*. 1,2,3. *Omarbeidet, forøget og forbedret Udgave*. 4. *Rettelser og tillegg ved Finn Hødnebø*. Kristiania/Oslo-Bergen-Tromsø

Grøtvedt, Per Nyquist. 1969, 1970, 1974. *Skrift og tale i mellomnorske diplomer fra Folden-området 1350-1450*. 1, 2, 3. Skrifter frå Norsk Målførearkiv ved Olav T. Beito. XXI, XXII, XXVIII. Oslo

Haugen, Einar. 1976. *The Scandinavian Languages. An Introduction to their History*. London

Hellevik, Alf et al. (eds.). 1966-. *Norsk Ordbok. Ordbok over det norske folkemålet og det nynorske skriftmålet*. Oslo

Holtsmark, Anne. 1951. Norsk språkhistorie. [Recension of Indrebø 1951.] *Nordisk tidskrift för vetenskap, konst och industri. Utgiven av Letterstedtska föreningen* 27 : 365-369

Huitfeldt, Henrik Jørgen (ed.). 1879. *Biskop Eysteins Jordebog (Den røde Bog)*. Christiania

Hægstad, Marius. 1899. Upphavet til det norske folkemaal. (Fyrste-fyrelesning av prof. M. Hægstad den 6/10 1899.) *Syn og Segn* 5 : 257-271

Hægstad, Marius. 1902. *Maalet i dei gamle norske kongebrev*. Videnskabsselskabets Skrifter. I. Historisk-filos. Klasse. 1902. No. 1. Kristiania

Hægstad, Marius. 1915, 1916, 1917. *Vestnorske maalføre fyre 1350*. 2 : 1, 2 : 2 : 1, 2 : 2 : 2. Videnskapsselskapets Skrifter. II. Hist.-filos. Klasse. 1914. No. 5, 1915. No. 3, 1916. No. 4. Kristiania

Indrebø, Gustav. 1951. *Norsk målsoga.* Ed. by Per Hovda and Per Thorson. Bergen

Johnsen, Oscar Albert. 1912. Ord- og sagregister. *Norges gamle Love. Anden række. 1388-1604. ... Første bind. 1388-1447. b. Registre.* Christiania

Kalkar, Otto. 1881-1885, 1886-1892, 1892-1901, 1902-1907, 1908-1918, 1976. *Ordbog til det ældre danske Sprog (1300-1700).* 1, 2, 3, 4, 5, 6. København

Keyser, Rudolf et al. (eds.). 1846-. *Norges gamle Love.* Christiania/Oslo. (Ref. to volume and number.)

Lasch, Agathe et al. (eds.). 1928-. *Mittelniederdeutsches Handwörterbuch.* Hamburg/Neumünster

Lind, Erik Henrik. 1905-1915. *Norsk-isländska dopnamn ock fingerade namn från medeltiden.* Uppsala/Leipzig

Lind, Erik Henrik. 1931. *Norsk-isländska dopnamn ock fingerade namn från medeltiden. Supplementband.* Oslo/Uppsala/København

Lübben, August. 1888. *Mittelniederdeutsches Handwörterbuch.* Norden/Leipzig. Repr. Darmstadt 1965

Munch, Peter Andreas (ed.). 1852. *Aslak Bolts Jordebog.* Christiania

Mæhlum, Brit. 1991. En oversikt over og en vurdering av kildegrunnlaget for sosiolingvistiske aspekter og innslag ved to perioder i norsk språkhistorieskrivning. *Maal og Minne* 1991 : 114-130

Pettersen, Egil. 1975. *Språkbrytning i Vest-Norge 1450-1550. Språket i vestnorske skrifter ved overgangen fra mellomalder til nyere tid.* Bergen-Oslo-Tromsø

Rindal, Magnus. 1987. *Brev frå Hamar 1350-1525. Gammalnorsk eller mellomnorsk språkform?* [Unpublished dissertation.] Bergen

Rindal, Magnus. 1993. Norsk språk 1350-1500. Gammalnorsk eller mellomnorsk? Ernst Håkon Jahr & Ove Lorentz (red./eds.). 1993. *Historisk språkvitenskap / Historical Linguistics.* Studier i norsk språkvitenskap / Studies in Norwegian Linguistics 5 : 395-404. Oslo

Ross, Hans. 1895, 1895-1913. *Norsk Ordbog.* Tillegg 1-6. Christiania. Repr. Oslo-Bergen-Tromsø 1971

Rygh, Oluf et al. (eds.). 1897-1936. *Norske Gaardnavne.* Kristiania/Oslo

Seip, Didrik Arup. 1934. *Studier i norsk språkhistorie.* Oslo

Seip, Didrik Arup. 1955. *Norsk språkhistorie til omkring 1370.* 2nd ed. Oslo

Simensen, Erik. 1992. Om ein plan for ei mellomnorsk ordbok – ordbok over skriftmålet i Noreg 1350-1550. Ruth V. Fjeld (ed.). 1992. *Nordiske studier i leksikografi. Rapport fra Konferanse om leksikografi i Norden 28.-31 mai 1991.* Skrifter utgitt av Nordisk forening for leksikografi. Skrift nr. 1 : 315-323. Oslo

Söderwall, Knut Fredrik. 1884-1918, 1891-1900, 1900-1918, 1953, 1953-1973. *Ordbok öfver svenska medeltids-språket.* 1, 2 : 1, 2 : 2, *Supplement.* Lund

Vroegmiddelnederlands Woordenboek. 5e concept-aflevering. (Leiden, 1993) W.J.J. Pijnenburg et al. (eds.)

Aasen, Ivar. 1873. *Norsk Ordbog med dansk forklaring.* 2nd ed. Christiania

The *Vroegmiddelnederlands Woordenboek* and the Computer

Tanneke Schoonheim

1. THE *VROEGMIDDELNEDERLANDS WOORDENBOEK*

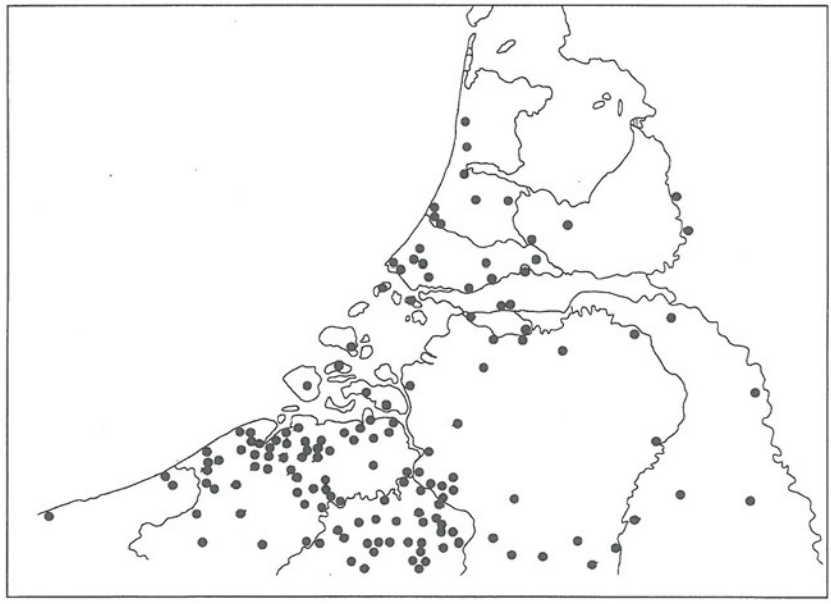
At the Institute for Dutch Lexicology in Leiden a team of four lexicographers has been compiling the *Vroegmiddelnederlands Woordenboek* (*Dictionary of Early Middle Dutch*) since 1988. This dictionary based on historical principles describes the vocabulary of the recorded thirteenth-century manuscripts written in the Middle Dutch dialects. The oldest literary products in Middle Dutch date from the beginning of this century and soon after the vernacular also came into use for administrational purposes. The language area involved stretches from Holland, Utrecht and the Lower Rhine in the North to Flanders, Brabant and Limburg in the South (see Figure 1). In 1998, the lexico-semantic description of the Early Middle Dutch word stock is to be completed.

2. SOURCES

The source material of the *Vroegmiddelnederlands Woordenboek* consists of three genres. First, there is the lexicographical material of the *Glossarium Bernense*, a Latin-Middle Dutch word list. Then, we have more than 2000 charters. Most of these are part of the first series of the so-called *Corpus*-Gysseling, an edition of almost all recorded official documents written in a Middle Dutch dialect up to and including the year 1300. Finally, we make use of almost all known literary texts and *artes* texts that have come down to us in thirteenth-century manuscript.

They, for the larger part, have been included in the second series of the *Corpus*-Gysseling. The total amount of word forms in the source material of the *Vroegmiddelnederlands Woordenboek* comes to 1,600,000. These will be described in approximately 28,000 different entries.

Figure 1. The Dutch-speaking area in the thirteenth century. The dots represent scribe centers. (W.J.J. Pijnenburg: "Het *Vroegmiddelnederlands Woordenboek (VMNW)*. Regionale en toevallige vormvarianten", in *Lexicologie. Een bundel opstellen voor F. de Tollenaere* (Groningen, 1977) pp. 209-220.



3. THE COMPUTER SYSTEM

The computer plays an important part in the realisation of the *Vroegmiddelnederlands Woordenboek*. Not only both the source material and the edited dictionary articles are stored in databases, but the analysis and interpretation of the source material take place in a computerised environment too. The lexicographical system used for the *Vroegmiddelnederlands Woordenboek* consists of three mutually linked

parts. The corpus database contains the complete source material for the dictionary. It is a relational database wherein all Early Middle Dutch word forms are provided with different kinds of additional information the lexicographer might need for the compilation of a dictionary article. The second part of the computer system is the so-called lexicographer's workbench. Here, the selected source material is shown to the lexicographer, so that he can interpret and analyse the data for the dictionary article. There are different options in which the lexicographer can choose to have the material presented to him. It can, for instance, be arranged alphabetically or chronologically, or with or without context. The third component of the computer system is the lexical database, also a relational one. There, the results of the lexicographer's analysis are put into an electronic form containing different fields of information: the elements of a dictionary article.

3.1. The corpus database

As said before, the corpus database with the source material of the *Vroegmiddelnederlands Woordenboek* is a relational database, containing approximately 1,600,000 word forms. As the *Corpus*-Gysseling was published at our institute, we had a version of most texts on tape so that it was relatively easy to transfer them into the computer. The few texts that were not available on tape were digitised through an optical reader. This resulted in a series of linear text files.

To prepare the material for lexicographical purposes all word forms were provided with a morphological code of three digits, referring to word class and inflection, and with a modern Dutch headword, enabling the lexicographer to gather all Middle Dutch spellings of a word (which can be quite a lot) at once. This process was done semi-automatically. Based on the indexes a valuation was made of the most probable code and headword for every word form type and these codes and headwords were automatically put into the texts. For example, all Middle Dutch word forms *ende*, mostly representing the conjunction *and*, but sometimes the noun *end*, were provided with the morphological code 800 (co-ordinated conjunction) and the headword EN (the modern Dutch equivalent of the conjunction). Hereafter, all texts were verified and the necessary corrections were entered manually. This meant that all instances where *ende* represented the noun *end* were changed from 800_EN to 000_EIND (see Figure 2).

Figure 2. A small piece of material from the linear text files containing documental information, Early Middle Dutch text, morphological codes and modern Dutch headwords.

```
<t an>
<n 0033>

<r Brug_22/8>
<d &1265>
(...)
<C 412_DAT> das <C 210+440_ZULLEN_MEN> salmen <C
250_GELDEN> ghelden <C 470+301_DE_EEN> dene <C
000_HELFT> helt
<L 008717> <C 700_TE> te <C 100_SINT> sent <C 022_BAVO>
baues <C 001_MIS> messe <C 800_EN> en#de$ <C
470+310_DE_ANDER> dander <C 000_HELFT> helt <C 700_TE>
te <C 000_HALF> half <C 021_MAART> marte.
```

After this, an automated verification procedure checked the more technical aspects of the text, for instance whether every word form had a code and a headword and whether every document had a place of origin and was properly dated. Finally, all texts were incorporated into the corpus database. The relational organisation of this database makes is easy to link the word form with information about, amongst others, headword, word class, date and place of origin, inflection and text genre (see Figure 3).

3.2. The lexicographer's workbench

The selected source material is copied into the lexicographer's workbench. Here, the lexicographer interprets and analyses the data, which can be presented in two ways. To get a first insight into the material he can ask the computer to give all word forms in combination with a summary of the most important information. It concerns information about the source, the document, page and line number in the text edition, genre, date and place of origin and a number that, amongst others, is important for the citations in the dictionary. This summary can be sorted alphabetically, chronologically, regionally or according to a combination of these. This enables the lexicographer to look at the variation in spelling and at the distribution in time, space and

Figure 3. The function "show detail".

DI Cterm 1

Commands Edit Customize Help

```
FICHE GEDETAILLEERD (RWO)                                        F01_121502
Woordvormnr: 30664  Fiche Kladblok
Redacteur  : TANNEKE      Act. Kode: 1  Omschr: bewijsplaatsanalyse
Ldb_lemma nr: 18891 Ldb_lemmatekst : maerte (I)
CG_lemma_nr : 112   CG-lemmatekst : MAART
Dokument Nr : 0033      Bronkode : 0  Genre : 0
Pagina/Regel/Woordvolgnr          : 0087  /12/211
Woordv.tekst : marte.
Morfkode   : 021  Morf. Omschr. : eigennaam, uitgang -e
Woordcombinatie nr:      Soort :   Deellemma :
Regio 1: 8   West-Vlaanderen                    Plaats 1  : 22  Brugge
Regio 2:                                        Plaats 2  :
Regiostatus : 1                  Kanselarij: 0
Datering    : 1265 - 1265
Tijdstatus  : 1                  Tijdsoort : 1
Redactioneel commentaar aanwezig: *
Literatuur advies aanwezig      : *
Toon Citaat                     : *

                        <F17> of <PF1>+<F17>= TERUG
```

genre. There are three more columns in this view I have not yet discussed; category and sub-category (to be seen as one item) and citation. With the help of the function of categorisation the lexicographer can classify his material and bring structure into it. In the citation column it is indicated whether the lexicographer has chosen a word form to be an illustrative example in the dictionary (see Figure 4).

For the determination of the exact meaning of a word, the word forms can also be presented in their context. The lexicographer can choose the size of the context, but normally the default concordance of 15 words before and 15 words behind the headword is used. There is little room left to show additional information directly on screen. Presented are headword, morphological code, categorial codes and citation indication (see Figure 5). If the lexicographer needs to see the rest of the information he can ask the computer for headword detail. With the help of these functions the lexicographer can analyse the meaning(s) and determine the inflection of the word he is describing. The material can be categorised more elaborately and citations can be formed on screen by marking off the beginning and the end of the Middle Dutch text surrounding the headword. When working with a lot of material it can sometimes be useful to have a printed version of the concordances. This contains both the concordances and the additional information from the summary.

Another tool for the description of large articles is the function that can sort the word forms according to their context. The lexicographer can define one or more word patterns and the computer will select only these cases answering to the search conditions. While working on the article *marte (I)* 'March', for instance, it became clear that this noun is often preceded by an adjective, indicating a more specified period of the month. The lexicographer can ask for all instances of the word *marte* directly preceded by an adjective (code 1**), as a result of which collocations as *inganghende / ingaende marte* ('the beginning of March'), *half / midde marte* ('half March') and *uteganghende / utegaende marte* ('the end of March') can very easily be distinguished (see Figure 6).

3.3. The lexical database

Finally, all information is gathered in the lexical database. The lexicographer uses an electronic form with different fields for the various information categories (see Figure 7). This way of storage makes it not

DECterm 1

Commands Edit Customize Help

TOON FICHE SAMENVATTING ldb_lemmanr : 18891

Sortering: jaar_tot,woordvormsort

rec#	bron	doku	pagina	regel	genre	datum	regio	plaats	wvorm_nr	cat	subcat	woordvormtekst
1	B	7000	248	03	W	1240	12		1066784	3	0	merte
2	308	2060	346	12	L	1253	7		979155	3	0	marte
3	0	0033	0087	17	A	1265	8	22	30664	2	9b2	marte.
4	0	0045	0103	51	A	1267	8	22	92184	1	9	marte
5	0	0046	0103	43	A	1267	8	22	92374	2	9b3	marte.
6	D	0019a	039	06	A	1271	6	7	28078	2	9a	maerte.
7	0	0118	0208	07	A	1272	8	22	44516	1	1	marte.
8	0	0131	0251	04	A	1273	8	22	110246	1	2	marte.
9	0	0147	0267	33	A	1274	8	22	49363	1	1	marte.
10	0	0148	0268	31	A	1274	8	22	110825	1	1	Marte.
11	0	0177	0299	40	A	1276	8	22	54681	1	3	m<A >ar</A>the
12	0	0207	0377	28	A	1277	8	22	134472	2	gb1	alfmarte
13	0	0220	0391	40	A	1278	8	22	62247	2	9b1	maerte.
14	0	0219	0391	02	A	1278	8	22	62055	1	3	marte.
15	0	0220	0392	02	A	1278	8	22	62325	1	3	marte.
16	0	0221	0392	34	A	1278	8	2	62489	2	9b1	marte
17	0	0266	0441	34	A	1279	8	22	69794			maerthe
18	0	0242	0420	33	A	1279	8	22	140097	1	1	m<A >ar</A>the.

[<PF1>+]<F17>terug <B>citaat <C>coll.cat. <D>detail <F>cat.decl. <G>goto <ENTER>(sub)cat. <F18>print <F20>store

Figure 4. The function "show summary".

Figure 5. The function "show concordances".

DECterm 1

Commands Edit Customize Help

FICHE CONCORDANTIES Ldb_lemmanr : 18891 Sortering op : JAAR_TOT + WOORDVORMTEKST F01_011519(PAS)

lemma_cg	morf	<---------- context ---------->	categorie	r	c
MAART	021t-ogen marcescere drogen marcidus droge marcjus	<C>merte mare see marinus uanden see maritimus an	3	0	Z J
MAART	021 </A> wine. bote<A >n</A> di des h't2ustes. De	<C>marte is so gedaen. du salt d's0ines l's0iues	3	0	Z J
MAART	021 baues messe en<A >de</A> dander helt te half	<C>marte. en<A >de</A> telken+dat dat vorse[i]de	2	9b2	Z J
MAART	021 vp den sat<A >er</A>dach vore meyedaghe. alf	<C>marte	1	9	Z N
MAART	021 rtins messe. en<A >de</A> dander helt te half	<C>marte. En<A >de</A> telken als dat vorseide la	2	9b3	Z J
MAART	021 rstine vten boghe .xxiiij. 't3d. nuwe te half	<C>maerte. ende ene hinne. 't% Jn Staecbruec hebw	2	9a	Z J
MAART	021 on .m'ho.cc'ho.lxx'ho. primo in die maent van	<C>marte.	1	1	Z N
MAART	021 +cc'ho+lxx'ho+ij'ho in onser vrouwen daghe in	<C>marte.	1	2	Z J
MAART	021 tseuentich+en<A >de</A>+drie jn die maent van	<C>marte.	1	1	Z N
MAART	021 zeuentich+en<A >de</A>+drie. jn die maent van	<C>Marte.	1	1	Z J
MAART	021 dit was ghedaen. tsfrindaghes na inganchende	<C>m<A >ar</A>the int iar ons heren als me<A >n</	1	3	Z J
MAART	021 . 't3s. helkes jaers te gheldene .v. 't3s. te	<C>alfmarte en<A >de</A> .v. 't3s. te si<A >n</A>	2	gb1	Z J
MAART	021 e helt temedewinter, ende dander helt te half	<C>maerte. den here Godeuarde, jof den ontvangher	2	9b1	Z J
MAART	021 euentich. vpten zesten dach van jngangh[en]de	<C>marte. Dit was Soendaghes na Grote+vastenauend	1	3	Z J
MAART	021 e+seuentich vpten zestendach bin inganghenden	<C>marte. Dit was Soendaghes na Grote+vastenauend	1	3	Z J
MAART	021 al, en<A >de</A> die and<A >er</A> elt te alf	<C>marte daer+na eerst comen sal, ende die ghelik	2	9b1	Z J
MAART	021 aertin. dien vorseiden volcarde. en<A >de</A>	<C>maerthe zinen wiue. al dit vorseide land. te h			Z N
MAART	021 on m'ho+cc'ho+lxx'ho+octauo jn die maent. van	<C>m<A >ar</A>the. vp die loue vande<A >n</A> ste	1	1	Z N

[<PF1>+]<F17>terug <F18>print <F20>store

<B>citaat <C>coll.cat. <D>detail <F>cat.decl. <G>goto <H>selectiegeg. <R>coll.red. <S>sorteer <PF1+S>=toon sleutels <ENTER>(sub)cat.

only possible to ask for Early Middle Dutch headwords or parts of headwords (for instance, all words ending with the suffix *-scap*), but also to search for, for instance, articles dealing with an adverb or a noun or articles containing only literary material or material from a certain document. What fields there are and the nature of the information they contain will be dealt with later on in this article.

The lexicographer can use special functions, which enable him to look through all the existing finished articles to see how certain problems have been solved earlier. This helps to maintain a certain degree of uniformity. Automatic procedures determine the place and date of the oldest occurrence, the total amount of occurrences and their frequency per genre. While filling in the form several checking procedures are active. For fields like *word sort* and *inflection*, for instance, the right abbreviation can be chosen from a table. It is also checked whether the fields that are essential for a dictionary article are filled in, for instance *headword, word sort* or *date and place of the oldest citation*. An article cannot be closed (that is, be given the status of finished) if one (or more) of these fields are left empty. The computer gives a warning and indicates what information is lacking.

3.4. The method of working

How is the actual practice of editing a dictionary article for the *Vroegmiddelnederlands Woordenboek*? The lexicographer works on a personal computer (which is linked to a mini computer on which all data is stored) with windows, containing the lexicographer's workbench and the lexical database (see Figure 8).

First, he registers a selected headword in the electronic form in the lexical database. Subsequently, the source material is selected in the lexicographer's workbench. Basically, these are the word forms carrying the chosen headword, but other word forms can also be added to the article.

When the lexicographer has selected all the necessary word forms, he can choose the way of presentation (as single word forms or in context), and the sorting order (alphabetically, chronologically, according to region or to a combination of these criteria). Based on this selection, the spelling of the Middle Dutch entry is chosen, whereby several basic rules have to be considered. The word sort is entered as well as the total amount of occurrences, the frequency per genre and the

Figure 6. The function "context sort".

DECterm 1

Commands Edit Customize Help

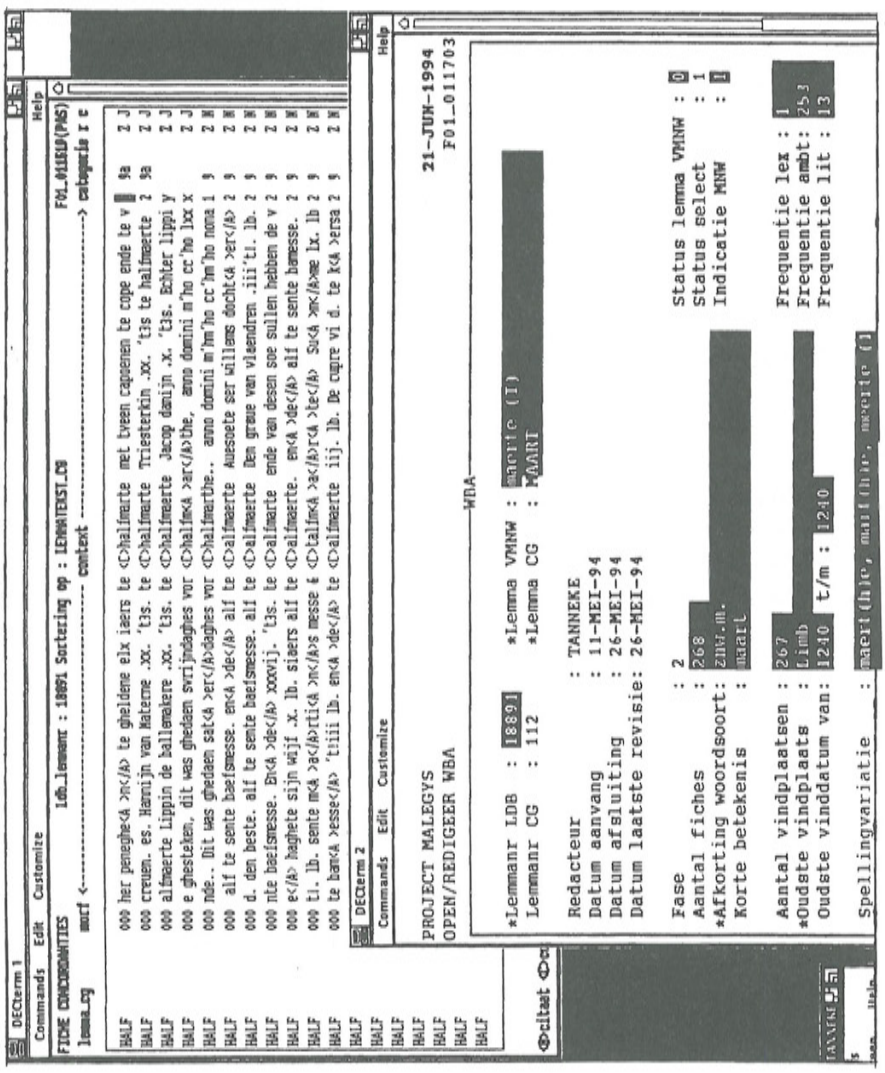
```
FICHES GESORTEERD OP CONTEXT     Ldb_lemmanr : 18891 Sortering op : LEMMATEKST_CG                F01_011519(PAS)
 lemma_cg       morf <---------------------------------------- context ----------------------------------------> categorie r c

INGAAND         101 s ghedaen tsat<A >er</A>daghes Na <S>jngaende <F>maerte.  Jnt jaer ons heren als men screef sijn  1  3     Z J
INGANGEND       101 t was ghedaen. tsfrindaghes na <S>inganchende <F>m<A >ar</A>the  int iar ons heren als me<A >n</  1  3     Z J
INGANGEND       101 ntich. vpten zesten dach van <S>jngangh[en]de <F>marte.  Dit was Soendaghes na Grote+vastenauend  1  3     Z J
INGANGEND       104 euentich vpten zestendach bin <S>inganghenden <F>marte.  Dit was Soendaghes na Grote+vastenauend  1  3     Z J
INGANGEND       101 r</A>dagh<A >es</A> na <S>ingangh<A >en</A>de <F>maerthe  anno domini m'ho cc'ho nonogesimo ----  1  3     Z N
INGANGEND       101 er des sonne<A >n</A>daghes na <S>inganghende <F>marthe  Ende dese brief was gegheue<A >n</A> in  1  3     Z N
INGANGEND       101 e bauesmesse & dander helt te <S>jngancghende <F>marthe  sonder coep & v<A >er</A>steruenesse. e  2  3b    Z J
INGANGEND       101 h<A >es</A> na <S>j<A >n</A>ganghend<A >e</A> <F>m<A >ar</A>the.  anno domini .M'ho. cc'ho. lxxx  1  3     Z N
INGANGEND       101 r des sonne<A >n</A>daghes na <S>inghanghende <F>marthe.                                        1  3     Z J
INGANGEND       101 en+En<A >de</A>+neghentich. te <S>Jnganghende <F>Maerte.                                        1  3     Z N
INGANGEND       101 h<A >es</A> na <S>j<A >n</A>ganghend<A >e</A> <F>m<A >ar</A>the.  anno domini .M'ho. CC'ho. Nona  1  3     Z N
INGANGEND       101 h<A >es</A> na <S>j<A >n</A>ganghend<A >e</A> <F>maerthe.  anno domini .M'ho.CC'ho. Nonagesimo.  1  3     Z N
INGANGEND       101 bi nachte sondagh<A >es</A> na <S>jnganghende <F>marte.  anno. domini .m'ho.cc'ho.lxxxxiiij. dit  1  3     Z N
INGANGEND       101 ne<A >n</A>+daghes. na <S>inghangen<A >de</A> <F>marte.                                         1  3     Z N
MIDDEN          101  lude van suethollant te gheldene te <S>midde <F>marte  die naest comet. Jn orconde van desen br  2  9a1   Z J
UITGAAND        101 ho.+cc'ho.xcvi. viii. daghe bin <S>vte+gaende <F>maerte.                                       1  6     Z J
UITGAAND        104  >ua</A>mmer+in achte daghe te <S>vte+gaenden <F>merte.  ende doe vant hi xlij mudde rogs. dese  1  6     Z J
UITGANGEND      101 it was ghedaen saterdaghes vor <S>vtganghende <F>marthe  anno domini m'hm'ho cc'hm'ho nonagesimo  1  6     Z J

                                  [<PF1>+]<F17>terug <F18>print <F20>store
<B>citaat <C>coll.cat. <D>detail <F>cat.decl. <G>goto <H>selectiegeg. <R>coll.red. <S>sorteer <PF1+S>=toon sleutels <ENTER>(sub)cat.
```

Figure 7. The first part of the electronical form of the entry *marte* (I) 'March'.

```
**************************************************
WBA: 18891     21-FEB-1995          16:15:02.22     Pagina 1
**************************************************
Lemmanr LDB:          18891
Lemma VMNW: marte (I)
Lemma sorteervorm:    marte (I)
Lemmanr CG:           112
Lemma CG:             MAART

Redacteur:            TANNEKE
Aanvang:              11-MAY-1994
Afsluiting:           26-MAY-1994
Revisie:              30-SEP-1994

Status lemma:         0
Status select:        1
Indicatie MNW:        1
Lemma MNW:            maerte

Fase:                 2
Aantal fiches:        268
Afkorting woordsrt:   znw.m.
Korte betekenis:      maart

Aantal vindplaatsen:  267
Oudste vindplaats:    Limb
Oudste datum vanaf:   1240
Oudste datum t/m:     1240

Frequentie lex:       1
Frequentie ambt: 253
Frequentie lit:       13

Spellingvariatie:     maert(h)e, mart(h)e, meerte (1x), merte
Morfologie:
Geo-info:             De vormen maerte, marte zijn uitsluitend
                      aangetroffen in Vlaanderen, Holland (7x,
```

	Dordrecht), Zeeland (2x) en de Nederrijn (1x). Limburg en West-Brabant kennen alleen de vormen meerte, merte. *De regio Brabant-Noord heeft tweemaal een vorm met* -a(e)- *en eenmaal een vorm met* -e-. *Opvallend is de vorm* merte *in Geraards bergen (1297).*
Etymologie:	< *Lat.* (mensis) martius *'(maand) maart'.*
Nb:	
Flexie soort:	ns
Flexie tekst:	-
Flexie soort:	ds
Flexie tekst:	-
Flexie soort:	as
Flexie tekst:	-
Nr onderscheiding:	I
Label:	
Overkoep. Betekenis:	
Nr betekenis:	1
Label:	
Definitie:	Maart, derde maand van het jaar.

Figure 8. The lexicographer's work bench and the lexical database in windows on screen.


DECterm 1
Commands Edit Customize Help
FIDRE CONCORDANTIES ldb_lemmanr : 18891 Sortering op : LEMMATEKST_CG F01_011703(PMS)
lemma_cg morf <---------- context ----------> categorie r c
HALF 000 her geneghe<A >K</A> te gheldene elx iaers te <D>halfmarte met tveen capenen te cope ende te v 9a Z J
HALF 000 creuen. es. Hannijn van Materne .xx. 'tis. te <D>halfmarte Triesterkin .xx. 'tis te halfmaerte 2 9a Z J
HALF 000 alfmaerte Lippin de ballemakere .xx. 'tis. te <D>halfmaerte Jacop danijn .x. 'tis. Echter lippi y Z J
HALF 000 e ghesteken, dit was ghedaen Swrijndaghes vor <D>halfmaerte ende van desen sue sullen hebben de v 2 9 Z M
HALF 000 nde... Dit was ghedaen sat<A >er</A>daghes vor <D>halfmarthe... anno domini m'hm'ho cc'hm'ho nona 1 9 Z M
HALF 000 alf te sente baefsmesse. en<A >dec</A> alf te <D>alfmaerte Auesaete ser willems dochte<A >er</A> 2 9 Z M
HALF 000 d. den beste. alf te sente baefsmesse. alf te <D>alfmaerte Den grene van vleendren .iii'tis. lb. 2 9 Z M
HALF 000 nde baefsmesse. En<A >de</A> xxxvij. 'tis. te <D>alfmarte ende van desen sue sullen hebben de v 2 9 Z M
HALF 000 e</A> haghete sijn wijf .x. lb. sieers alf te <D>alfmaerte. en<A >de</A> alf te sente bamesse. 2 9 Z M
HALF 000 ti. lb. sente m<A >a</A>rtin<A >K</A>s messe & <D>talfm<A >a</A>rt<A >te</A> Su<A >K</A>se lx. lb 2 9 Z M
HALF 000 te ban<A >esse</A> 'tiiii lb. en<A >de</A> te <D>alfmaerte iij. lb. De cupre vi d. te k<A >ersa 2 9 Z M
HALF
HALF
HALF
HALF
HALF
HALF
<Select <Cr
DECterm 2
Commands Edit Customize Help
PROJECT MALEGYS 21-JUN-1994
OPEN/REDIGEER WBA F01_011703
WBA
*Lemmanr LDB : 18891 *Lemma VMNW : maerte (I)
Lemmanr CG : 112 *Lemma CG : MAART
Redacteur : TANNEKE
Datum aanvang : 11-MEI-94
Datum afsluiting : 26-MEI-94
Datum laatste revisie: 26-MEI-94
Fase : 2 Status lemma VMNW : 0
Aantal fiches : 268 Status select : 1
*Afkorting woordsoort: ZNW.mh. Indicatie MNW : 1
Korte betekenis : maarl
Aantal vindplaatsen : 267 Frequentie lex : 1
*Oudste vindplaats : Limb Frequentie ambt: 253
Oudste vinddatum van: 1240 t/m : 1240 Frequentie lit : 13
Spellingvariatie : maert(hie, maerthie, maerte (I)
TANNEKE
Help

place and date of the oldest occurrence. If relevant, a link is made to the *Middelnederlandsch Woordenboek* (*Middle Dutch Dictionary*), describing the Middle Dutch stock of words from 1280 till 1550. Further, there are fields for spelling variation, morphology, etymology and dialect-geographical information.

To determine the inflection and the meaning of the word the lexicographer uses the concordances in the lexicographer's workbench. He examines all occurrences and orders them according to their meaning. The most representative occurrences are used as citations. Hereafter, the structure of the article is put into the lexical database by giving every meaning a number and definition and by indicating collocations and special means of usage. A survey of the complete structure of the article is given in the definition in short. Then, the citations can be transferred automatically from the lexicographer's workbench to the lexical database by means of the word form number (when dealing with only one citation at a time) or the categorial code (when dealing with more citations at the same time). The advantage of this method is not only that the text of the citations is copied faultless into the lexical database, but that all information about the source, page and line number and date and place of origin is copied as well. This does not only diminish the amount of typing work considerably, it also saves much correction work (see Figure 9).

Some fields in the lexical database are for internal use only. Examples are the name of the lexicographer who made the article, the date on which the article was begun, the date it was finished and an indication of the reference works. Based on the latter, for instance, the bibliography can be generated automatically. Finally, the article is ready and can be stored definitively in the lexical database.

There are two ways to print the result of the lexicographer's work. First, it is possible to print the article in database structure, which is useful when checking whether all fields contain the correct information. It is also possible to print the article as it will appear in the dictionary. For this we make use of two conversion programmes. The first programme translates the information from the different fields in the database to a sequential file, which is structured by various lay out codes. The other programme performs the conversion to WordPerfect and delivers a file that looks like a real dictionary article (see Figure 10). Now the lexicographer can see if the information from the various fields forms a nice, clear and conveniently structured article. If not, he

Figure 9. An Early Middle Dutch citation with all necessary information copied alongside. Text codes will be converted to WordPerfect characters later on, for instance, 's0i becomes í.

CITAAT (buiten verbinding):	
Nr citaat:	6
Tekst citaat:	De marte is so gedaen. du salt d's0ines l's0iues grote hude hebben dan.
Vindplaats: Gez.reg.	
Herkomst:	Nrijn
Nr dokument:	2060
Blz van:	346 t/m:
Regel van:	12 t/m:
Jaar van:	1253 t/m: 1254

can reopen the entry and make the necessary changes.

4. CONCLUSION

The advantage of making a dictionary with the help of a computer lies not in the domain of the actual interpretation and analysis of the data. The gain, however, is to be found in the amount of time needed for the preliminary work: the selection and presentation of the material. Moreover, as the computer takes over some of the checking procedures, a higher degree of uniformity is guaranteed and less time is needed for correction. Another advantage is the flexible method of working. Whereas conventional dictionaries often have to work alphabetically, we can work thematically and describe all words referring to the same subject at once, irrespective of their alphabetical place in the dictionary. Because of the fact that the dictionary will not be published before the end of the project in 1998, rectifications can be made and new insights and material can systematically be added to all articles. The general conclusion may be that the use of a computer does not automatically result in a better dictionary, but that it has certain advantages in enabling one to give extra attention to the lexicographer's part of the work: the interpretation and analysis of the data.

Figure 10. A part of the entry *marte (I)* 'March' in the *Vroegmiddelnederlands Woordenboek.*

marte (I) znw.m.

maart [267]

(maerte) 1/253/13 Limb 1240. maert(h)e, mart(h)e, meerte *(1x)*, merte; *ns* -; *ds* -; *as* - < *Lat.* (mensis) martius *'(maand) maart'. De vormen* maerte, marte *zijn uitsluitend aangetroffen in Vlaanderen, Holland (7x, Dordrecht), Zeeland (2x) en de Nederrijn (1x). Limburg en West-Brabant kennen alleen de vormen* meerte, merte. *De regio Brabant-Noord heeft 2x een vorm met* -a(e)- *en 1x een vorm met* -e-. *Opvallend is de vorm* merte *in Geraardsbergen (1297).*

I 1. Maart, derde maand van het jaar. *De Romeinse tijdrekening begon oorspr. in maart, vanaf circa 153 v.C. werd het echter gebruik om het nieuwe jaar in januari te laten beginnen (Strubbe/Voet, p. 51, 59)* ‖ marcjus : merte, *Bern. 248,3 (Limb 1240).* De marte is so gedaen. du salt dínes líues grote hude hebben dan, *Gez.reg. 346,12 (Nrijn 1253).*

1.1. Maart, als deel van een datering. ‖ Dese pointinghe ghinc in int iaer van lxxxi vpden eersten dach van maerte, *Corp.I 613,34-35 (Roeselare? 1282).*

(...)

In de verb. ingaende/inganghende marte; die inganc/dat ingaen van marte. **Begin maart, mog. meer bep.: 1 maart.** ‖ dit was ghedaen. tsfrindaghes na inganchende m*ar*the, *Corp.I 299,39-40 (Brugge 1276).*Dit was ghedaen vp den jnganc van marthe. *Corp.I 1190,39-40 (Brugge 1287).* tsat*er*daghes Na jngaende maerte. *Corp.I 1430,34 (Brugge 1290).* int jngaen van marte, *Corp.I 1858,28 (Bergen-op-Zoom/Drongen 1294).* Smaendaghes. ten jnghanghe van marte. *Corp.I 317,31 (Oudenaarde 1312).*

In de verb. half marte. **Half maart, mog. meer bep.: 16 maart.** ‖ Dese lettren waren ghemaect ende ghegheuen jnt van den jncarnatione hons heren ih*e*sus kerst .M°.CC°. ende .lxxx°. iar omtrent alf. marte. *Corp.I 551,40-42 (Gent 1281).*

(...)

In de verb. onser vrouwen dach/avont in marte. **Onze-Lieve-Vrouwedag in maart, Maria-Boodschap, 25 maart.** ‖ in onser vrouwen daghe in marte. *Corp.I 251,3-4 (Brugge 1273).*

1.2. Maart, als periode voor betaling van pacht of huur.
1.2.1. Als betalingsperiode van een jaarlijkse pacht of huur. ‖ Echt*er* van ons*er* m*er*sch te Berchem die ionfrou*en* katelin*en* va*n* grieq*uart* was .v. õ. te maerte. *Corp.I 1080,32-33 (Oudenaarde 1285).*
(...)
In de verb. te midde marte. **Half maart, mog. meer bep. de 16e.** *Slechts 1x aangetroffen.* ‖ *(Geoorkond wordt)* dat wi onse porte van dordrecht gheloven met d[e]sen brieve quite ende scadeloes te houden. van dien ses hondert ponden ende dertich daer zi borghen af ziin ieghens haren janne pac van brucghe voer onse lude van suethollant te gheldene te midde marte die naest comet. *Corp.I 1362,17-21 (Holland, graf.kans. 1289).*
1.2.2. Als betalingsperiode voor een halfjaarlijkse pacht of huur.
In de verb. te inganghende marte. **Begin maart, mog. meer bep.: op 1 maart.** *De termijnen zijn 1 oktober en 1 maart.* ‖ o*m*me tiene schelleghe tsiars goeder vlams*cher* peneghe*n* tegheldene elkes jars. dene helt sinte bauesmesse & dander helt te jngancghende marthe, *Corp.I 787,37-39 (Brugge 1284).*
(...)

References

Corpus van Middelnederlandse teksten (tot en met het jaar 1300); reeks I: ambtelijke bescheiden. Uitgegeven door M. Gysseling. 's-Gravenhage 1977. 9 vols.

Corpus van Middelnederlandse teksten (tot en met het jaar 1300); reeks II: literaire handschriften. Uitgegeven door M. Gysseling. 's-Gravenhage, 1980 - Leiden, 1987. 6 vols. (II-1 Fragmenten; II-2 Jacob van Maerlant: *Der Naturen Bloeme*; II-3,4 Jacob van Maerlant: *Rijmbijbel*; II-5 *Sente Lutgart*; II-6 *Sinte Lutgart*; *Sinte Kerstine*; *Nederrijns Moraalboek*)

Het Glossarium Bernense. Een Vroegmiddelnederlandse tweetalige Latijns-Limburgse woordenlijst. Uitgegeven door L. de Man en P.G.J. van Sterkenburg. 's-Gravenhage 1977.

Het Luikse Diatessaron. Uitgegeven door C.C. de Bruin. Met de Engelse vertaling van A.J. Barnouw. Leiden, 1970.

Middelnederlandsch Woordenboek door E. Verwijs en J. Verdam. 9 vols. 's-Gravenhage 1885-1929.

Het oudste goederenregister van Oudenbiezen (1280-1344). [Uitgegeven] door J. Buntinx en M. Gysseling. Dl. 1, tekst. Tongeren, 1965.

The Importance of Legal History for Old Frisian Lexicography

Oebele Vries

In this paper I shall explain why knowledge of legal history is of crucial importance for a lexicographer of Old Frisian and how use can be made of it most efficiently for a projected dictionary (glossary) of the language of the Old Frisian charters. But first of all I must briefly introduce Old Frisian and the state of the art of Old Frisian lexicography.

The periodisation of Frisian is rather confusing, not being parallel to that of the other West Germanic languages (Sjölin 1969:6-8; Markey 1981:40-45). The boundary of Old Frisian is set as late as 1550, when Frisian, be it temporarily, ceased to be one of the written languages of Friesland. There is another peculiarity of Old Frisian that deserves attention, the fact that our knowledge of this language is based on extremely one-sided material (cf. Gerbenzon 1958:3; Munske 1973:2). Almost the whole corpus of Old Frisian texts that have survived consists of legal texts (laws, charters, records of juridical proceedings); amazingly few literary and religious texts are extant and only a small number of chronicles and private letters have come down to us (survey of Old Frisian texts: Sjölin 1969:9-10; Markey 1981:68-100). This largely explains why the field of Old Frisian studies, especially as far as text edition and lexicography are concerned, has been dominated for a considerable time by legal historians, one of the first and undoubtedly the most outstanding being Karl Freiherr von Richthofen. At the instigation of nobody less than Jacob Grimm this German nobleman erected a monument in the field of Germanic studies by publishing, as early as 1840, an edition of all Old Frisian legal texts that were accessible to him

at the time (Richthofen 1840a) as well as a comprehensive dictionary of Old Frisian (Richthofen 1840b).

By now Von Richthofen's text edition has been completely replaced by modern, mostly diplomatic editions, but up to now his dictionary, antiquated as it is, has not been superseded. So far two projects to compile a new comprehensive dictionary of Old Frisian, comparable in scope to Von Richthofen's, have been launched: at the Universities of Hamburg in 1939 and of Groningen in 1948 respectively. The Hamburg initiative was aborted as a result of the Second World War, even before it had started properly (Bachofer and Beck 1991:671). The Frisian Institute of Groningen University, which assumed responsibility for the Old Frisian dictionary project, has been more successful in preparing text editions, most of them provided with glossaries, as an indispensable prerequisite for compiling a new Von Richthofen, than in carrying out the task of compiling the dictionary itself. Up to now its main performance in the field of lexicography has been the processing of the language material of all Old Frisian text editions on partly hand-written, partly mimeographed, word slips. Each word slip contains a text fragment, one word of which is marked. The words thus marked have subsequently been alphabetised and lemmatised (Sjölin 1972-3). In 1985, as a consequence of the introduction of a new system of research funding by Dutch universities, our institute had to set new research priorities, which left almost no time for the Old Frisian dictionary project. In this situation the decision was taken to limit the scope of the project by focusing on compiling a corpus dictionary based on the Old Frisian charters (about 1100), and similar texts, that is, documents of juridical and administrative proceedings (several hundreds of pages), and letters (about 100). This represents a segment of the Old Frisian language tradition uncovered by Von Richthofen, to whom these sources, with the exception of a small number of charters, had not yet been available. The projected charter dictionary should be a comprehensive and not a concise one, as a concise dictionary of all Old Frisian, the language of the charters and similar texts included, was being compiled by Hofmann (Hofmann 1989:242-259).

An important consideration in making the afore-mentioned decision was the fact that all Old Frisian charters and similar texts have now been edited, but, with one, relatively unimportant exception, with no glossaries at all. An additional consideration was the announced publication of a comprehensive Dictionary of Middle High German

Charters (*Wörterbuch der Mittelhochdeutschen Urkundensprache*), the first issue of which appeared in 1986. This German dictionary not only covers a corpus that is entirely comparable to the one on which our choice had fallen (general introduction to charters, especially in German: Kirschstein 1984), but it is also structured in a way that could serve as an example for our own project (see also Dittmer 1990; Schulze 1990).

Our intention to compile a comprehensive dictionary of the Old Frisian charters and similar texts was made public in an article published in 1990 (Vries 1990a). It was accompanied by a four page sample entry of the verb *biliva*, i.e. to remain (German *bleiben*). When this paper was read (1994), it was not yet clear whether the Frisian Institute of Groningen University would be able to carry out this project (see for the state of the art of Old Frisian lexicography at that time, Århammar 1990, section 2; Bremmer 1994, section 1). In 1996, however, an agreement was signed by the University of Groningen and the Fryske Akademy (Ljouwert/Leeuwarden) concerning the continuation of the Old Frisian dictionary project under mutual responsibility. It was determined, moreover, to return to the original idea of compiling a comprehensive dictionary of all Old Frisian. As Hofmann had indicated shortly before he became unable to complete the compilation of his concise dictionary, at least in the intended form, it was decided that the projected charter dictionary should take the form of a glossary. The preparatory work on this glossary was taken up in 1997.

Being a historian, my interests lie in the amount of research necessary into the historical backgrounds of the period and region of a medieval language described in a dictionary. I want to emphasise that for the lexicographer of Old Frisian, most Old Frisian texts being legal documents, a thorough knowledge of especially legal history is obligatory, because without this knowledge many errors and inaccuracies in the field of semantics can be expected (with regard to charters cf. Schubert 1979:11). This does not mean, however, that charters and other legal documents include law terms only or that the intended Old Frisian Charter Dictionary will be a dictionary exclusively of law terms, like the *Deutsches Rechtswörterbuch* (German Law Dictionary). The language of the charters, as has been emphasised by several authors (e.g. Drüppel 1984:2), consists of technical juridical as well as everyday language (and, I would add, stereotypical charter formulas). In the Old Frisian charters, too, we encounter many terms from everyday life, and of course semantic problems can arise from

these as well, for instance in determining the exact meaning of designations of garments or jewellery, which are frequently mentioned in wills. In this paper I concentrate on law terms, however. As far as this category of terms is concerned, the still unfinished *Deutsches Rechtswörterbuch*, the first issue of which was published in 1914 (cf. Schmidt-Wiegand 1990), is more often than not a very useful tool for solving semantic problems. A proposal to compile a comparable historical dictionary of Dutch juridical language was launched in 1941 (Ganshof 1941), but this plan has never materialised. As far as Dutch is concerned we have to make do with Stallaert's glossary of outdated legal terms and words from everyday life, collected from sources from the southern part of the Netherlands (Flanders, Brabant and Limburg) only. This work has many deficiencies, but it can still be useful to us. Recently a new project to compile an historical dictionary of Dutch legal terminology has been launched, however.

It is obvious that much can be gained in the field of semantics if, before starting to write the entries of the projected dictionary of the language of the Old Frisian charters and similar texts, legal terminology as used in this corpus is studied more closely than has been the case up till now. With regard to this it is important to know that there is a considerable difference between the legal terms of the earlier Old Frisian law manuscripts (up till the fourteenth century), which as a rule have been explained sufficiently, and those from the charters and similar texts, which originated in a period as late as 1370-1550. For instance, a word that is well known from older sources sometimes has a more specialised or even a quite different meaning in the context of the charters (cf. Schmidt-Wiegand 1993a:598-9). More often, however, we encounter new terms, frequently abstract nouns and newly formed compounds. In order to explain this emergence of a new terminology we should bear in mind that the vernacular charter was a quite new text category, which developed relatively late, especially in the case of Old Frisian. The point is that new, often complex matters of a juridical or economic nature had to be expressed in the vernacular, in our case in Old Frisian, which entailed the need to extend its vocabulary considerably (cf. De Boor and Haacke 1957:XXII). Moreover, in the charters and records of juridical proceedings canon and Roman law are heavily represented, whereas the older Frisian law manuscripts reflect traditional Germanic law. I shall return to this matter below.

In order to come to terms with the specific terminology of the language of the Old Frisian charters and similar texts it is necessary to gain an insight into the *origin* of this terminology. An important point in this respect is the fact that proof by charter was accepted in Friesland hesitatingly and quite late, i.e. in the early part of the fourteenth century, much later than in the neighbouring county of Holland, for example. Moreover, as has already been said, Latin as language of the charters was substituted by the vernacular in the Frisian territories at a much later stage than in the adjacent Dutch speaking areas, about 1370. As in Friesland, not only Frisian came into use as a charter language, but, be it to a much lesser extent, Dutch as well, we may assume a strong influence of Dutch charter terminology on its Frisian counterpart. Currently, general studies of Old Frisian as well as Middle Dutch charter terminology are still lacking, so we cannot draw any easy conclusions. As far as Old Frisian is concerned we can make use, however, of a small, but important study by Gerbenzon, a legal historian, called *Friese rechtstaal en vreemd recht* (Frisian legal terminology and foreign law) (Gerbenzon 1958).

Gerbenzon sketches the influence of the infiltration and reception of canon and Roman law in Friesland on Old Frisian legal terminology. According to this author, in coining words for concepts derived from Roman law a choice was made between two possibilities. The first was to make use of the existing Old Frisian vocabulary by giving old words a new meaning, the second was to resort to loan-words directly derived from Latin, more often than not slightly modified. Gerbenzon shows that as a rule till the fifteenth century the former method was chosen, but after that time the latter dominated. He demonstrates this phenomenon mainly on the basis of law books, deriving examples from one category of charters only: wills. The instances given by him are strictly limited to legal terms from Roman law. From these examples we gather the impression that, contrary to our expectation, there was almost no influence of Middle Dutch on Frisian legal terminology. This does not imply, however, that everyday language and stereotypical charter formulas, as used in the Old Frisian charters, were not influenced by Middle Dutch either. On the contrary, without going into this more closely, it can be stated that the opposite is true.

My contention is that legal terms in the context of the Old Frisian charters and similar texts can most efficiently be analysed by not studying them in isolation (i.e. simply according to their alphabetical

order, which is normally the case when dictionary entries are written), but according to what I call *legal word fields*. This can be achieved by splitting up the corpus of charters according to their juridical subject matter, subsequently isolating the terms that are typical of each legal word field, and finally comparing these terms with the terminology of similar word fields in Latin as well as in adjacent vernaculars (in our case Middle Dutch and Middle Low German). This method of comparative investigation of legal vocabulary is related to, but not identical with, what in German is called *Rechtswortgeographie* (legal term geography), a direction of philological investigation which studies the geographical distribution of legal terms (Munske 1968). The starting point in this method is the legal term, most of the time a so-called "inherited word" (German: *Erbwort*). As far as Old Frisian is concerned we can boast one fundamental work in this field: Part II of Munske's impressive monograph on the legal vocabulary in the domain of crime in the earlier West Germanic legal sources (Munske 1973:79-210). As this vocabulary is limited to the older sources (ninth to fourteenth centuries), we can, however, hardly make use of this work when studying the language of the Old Frisian charters.

I myself have applied the legal word field method to the terminology of the arbitration procedure in all Old Frisian charters that are concerned with this matter, 113 in all, and in a paragraph of an Old Frisian law book from the late fifteenth century (Vries 1992; see, for a comparable application of this method: Vries 1996). The arbitration procedure, which is based on Roman and canon law, is very suitable for comparative investigation for three reasons: this procedure was very frequent in late medieval Friesland, it was also well known in the adjoining regions, and it has a very rich and characteristic terminology.

The results of this kind of approach are twofold. First of all, synonyms can more easily be recognised and small, but sometimes important differences in meaning between apparent synonyms can be worked out more efficiently than by studying these words (or expressions) in isolation from each other, simply in their alphabetical order. Secondly, it shows to what extent a language coined its own terminology and how far it borrowed from other vernaculars. It could be established, for instance, that a considerable part of Old Frisian terminology in the field of the arbitration procedure, contrary to Gerbenzon's conclusions, can be found in Middle Dutch and/or Middle Low German as well. Another part of Old Frisian terminology, however, can be said

to consist not of loan translations from Middle Dutch or – maybe – Middle Low German, but of existing Old Frisian words that had acquired a new meaning, or of independently formed new words.

Let me give some examples of terms in the field of the arbitration procedure that can be said to be uniquely Old Frisian:

baer	literal meaning: what is public; denotes in older sources: charge before a law court; in arbitration terminology: verdict (of arbitrators);
eg, ig	literal meaning: edge; denotes in arbitration terminology: party (in a dispute).

So in these two cases we have to do with existing Old Frisian words that were given a new meaning. In the former case, moreover, this even seems to have pushed aside the earlier one completely. Another word that acquired a new meaning is almost certainly the following. We cannot prove this, however, as, contrary to the first two, it is not attested in older texts:

skiwa	literal meaning: to split; denotes in arbitration terminology: to separate (the disputing parties).

Most likely a newly formed abstract noun is:

skortinge	literal meaning: difference; denotes: dispute.

Undoubtedly a newly formed word is:

ura	a substantivisation of the preposition *ur* (over); denotes: chief arbitrator (known in Middle Dutch and Middle Low German as *overman*). This term is also quite interesting from a morphological point of view as substantivisations of this type are extremely rare.

An important change in Old Frisian arbitration terminology was brought about by the massive penetration of terms that were directly derived from Latin, a development which can be observed after about 1500. In the fifteenth century some slightly adjusted Latin terms were already in use in the field of the arbitration procedure, as our corpus

shows. The most striking example is *compromis*, literally 'compromise', denoting the submission to arbitration of a dispute by the opposing parties. It comes as a surprise that this is the sole word we encounter in Old Frisian texts to denote this important concept. However, to my knowledge, the same holds for Middle Dutch and Middle Low German.

After 1500 arbitration terms derived directly from Latin abound in Old Frisian, more and more frequently ousting the Frisian equivalents that had been in use till then. Examples are: *questie* (first attestation in 1521) for denoting the concept 'dispute'; *part* (first attestation about 1480) or *partie* (first attestation in 1535) for denoting the concept 'party in a dispute'; *examineria* (first and only attestation in 1506) for denoting the concept 'to examine', i.e. the documents handed in by the disputing parties. A rather peculiar derivative of the term *compromis*, which denotes, as we have seen, the submission to arbitration of a dispute, is the verb *bicompromissia*, which occurs, just like the word it was derived from, at a rather early time (first attestation in 1441). Interestingly enough, at about 1500 this form was eliminated by a more correct one, at least from the view of Latin morphology, i.e. *bicompromitteria*. These and similar words derived directly from Latin were integrated into the vernacular by becoming linked to well-known Old Frisian terms with the same meaning, in so-called 'pair formulas', e.g. *skortinge ende questie*, both terms meaning 'dispute'; and *parten ende iggen*, both terms meaning 'parties in a dispute' (cf. Schmidt-Wiegand 1993b:1017). By this the 'pair formula', which occupied an important place in earlier Germanic legal language, acquired a quite new meaning, far away from its original use (cf. Dilcher 1961:76).

How can the penetration of a host of Latin terms in the vernacular at such a late stage be explained? The answer to this question has already been given by Gerbenzon (1958:16-17): in 1499, after the Frisians had lost their freedom from sovereign rule, their new lord, Duke Albrecht of Saxony, founded a High Court, which was composed, at least after a short time, mostly of jurists versed in Roman law, many of them being, in addition, non-Frisians. The result of this was that the scribes of charters, not only in the field of arbitration, but in general (at least if they were familiar with Roman law themselves) started using terms and expressions derived directly from Latin, in order to avoid the risk that charters written by them, which might at some time or another play a role in a lawsuit before the High Court, would not be properly understood by the judges of this distinguished institution. Another

consequence of this new situation was a gradual switch to Dutch, the language used by the chancery of the High Court, as the language in which charters were written, public as well as private ones. This turn was so radical that after 50 years Frisian had completely been pushed aside as a charter language by Dutch (Vries 1990b; Vries 1993).

To my knowledge, investigations of legal word fields, as encountered in the charters, are still lacking in other languages, as earlier studies of charter vocabulary concentrate on stereotypical charter formulas, not on legal terminology (e.g. Cordes 1959, Matzinger-Pfister 1972, De Boor 1975). I hope to have shown that for compiling a dictionary (or glossary) based on Old Frisian charters and similar texts, that not only is a profound knowledge of legal history necessary, but that such a project can take advantage of a thorough study of legal terminology according to legal word fields. I assume that the same could be the case with similar dictionary projects of medieval languages other than Old Frisian.

References

Århammar, Nils 1990. "Friesische Lexikographie". In: *Wörterbücher/Dictionaries/Dictionnaires. Ein internationales Handbuch zur Lexikographie*, vol. II. Eds. Franz Josef Hausmann, Oskar Reichmann, Herbert Ernst Wiegand and Ladislav Zgusta. Berlin-New York (2022-2036).

Bachofer, Wolfgang, and Wolfgang Beck 1991. "Deutsche und Niederdeutsche Philologie. Das Germanische Seminar zwischen 1933 und 1945". In: *Hochschulalltag im <<Dritten Reich>>. Die Hamburger Universität 1933-1945*, vol. II. Eds. Eckart Krause, Ludwig Huber and Holger Fischer. Berlin-Hamburg (641-703).

Boor, Helmut de 1975. *Actum et Datum. Eine Untersuchung zur Formelsprache der deutschen Urkunden im 13. Jahrhundert.* München.

Boor, Helmut de, and Diether Haacke 1957. "Vorrede". In: *Corpus der altdeutschen Originalurkunden bis zum Jahr 1300*, vol. III. Eds. Helmut de Boor and Diether Haacke. Lahr (VII-XLI).

Bremmer Jr, R.H. 1994. "Frisian Lexicography". In: *The Encyclopedia of Language and Linguistics*, vol. III. Eds. R.E. Asher and J.M.Y. Simpson. Oxford-New York-Seoul-Tokyo (1309-1310).

Cordes, Gerhard 1959. "Zur Erforschung der Urkundensprache". In: *Jahrbuch des Vereins für niederdeutsche Sprachforschung* 82:63-79.

Deutsches Rechtswörterbuch (Wörterbuch der älteren deutschen Rechtssprache), vols. I-... Weimar, 1914-....

Dilcher, Gerhard 1961. *Paarformeln in der Rechtssprache des frühen Mittelalters*. Darmstadt.

Dittmer, Ernst 1990. "Über das Wörterbuch der mittelhochdeutschen Urkundensprache". In: Schützeichel and Seidensticker 1990:46-58.

Drüppel, Christoph Josef 1984. *Altfranzösische Urkunden und Lexikologie. Ein quellenkritischer Beitrag zum Wortschatz des frühen 13. Jahrhunderts*. Beihefte zur Zeitschrift für romanische Philologie 203. Tübingen.

Ganshof, F.L. 1941. *Voorstel tot voorbereiding en uitgave van een historisch woordenboek der Nederlandsche rechtstaal* (met een verslag door E.I. Strubbe). Mededeelingen van de Koninklijke Vlaamsche Academie voor Wetenschappen, Letteren en Schoone

Kunsten van België. Klasse der letteren en der moreele en staatkundige wetenschappen, Jaargang III, N° 3. Brussel.

Gerbenzon, P. 1958. *Friese rechtstaal en vreemd recht.* Groningen.

Hofmann, Dietrich 1989. *Gesammelte Schriften*, vol. II: *Studien zur Friesischen und Niederdeutschen Philologie*. Hamburg.

Kirschstein, Bettina 1984. "Urkunden". In: *Reallexikon der deutschen Literaturgeschichte*, vol. IV. Eds. Klaus Kanzog and Achim Masser. Berlin-New York (592-603).

Markey, T.L. 1981. *Frisian.* Trends in Linguistics. State of the Art Reports 13. The Hague-Paris-New York.

Matzinger-Pfister, Regula 1972. *Paarformel, Synonymik und zweisprachiges Wortpaar. Zur mehrgliedrigen Ausdrucksweise der mittelalterlichen Urkundensprache*. Zürich.

Munske, Horst Haider 1968. "Rechtswortgeographie". In: *Wortgeographie und Gesellschaft*. Ed. Walther Mitzka. Berlin (349-370).

Munske, Horst Haider 1973. *Der germanische Rechtswortschatz im Bereich der Missetaten. Philologische und sprachgeographische Untersuchungen, I. Die Terminologie der älteren westgermanischen Rechtsquellen*. Berlin-New York.

Richthofen, Karl von 1840a. *Friesische Rechtsquellen*. Berlin (reprint: Aalen 1960).

Richthofen, Karl von 1840b. *Altfriesisches Wörterbuch.* Göttingen (reprint: Aalen 1961).

Schmidt-Wiegand, Ruth 1990. "Das 'Deutsche Rechtswörterbuch', Geschichte und Struktur". In: Schützeichel and Seidensticker 1990:155-168.

Schmidt-Wiegand, Ruth 1993a. "Urkundensprache". In: *Handwörterbuch zur deutschen Rechtsgeschichte*, 35. Lieferung. Berlin (593-602).

Schmidt-Wiegand, Ruth 1993b. "Volkssprache". In: *Handwörterbuch zur deutschen Rechtsgeschichte*, 36. Lieferung. Berlin (1010-1021).

Schubert, Fritz 1979. *Sprachstruktur und Rechtsfunktion. Untersuchung zur deutschsprachigen Urkunde des 13. Jahrhunderts*. Göppinger Arbeiten zur Germanistik 251. Göppingen.

Schulze, Ursula 1990. "Anmerkungen zum 'Wörterbuch der mittelhochdeutschen Urkundensprache' aus der Perspektive der Herausgeber und Bearbeiter". In: Schützeichel and Seidensticker 1990:189-197.

Schützeichel, Rudolf, and Peter Seidensticker (Eds.) 1990. *Wörter und Namen. Aktuelle Lexikographie. Symposium Schloß Rauischholzhausen 25.-27.9.1987*. Marburg.

Sjölin, Bo 1969. *Einführung in das Friesische*. Sammlung Metzler M 86. Stuttgart.

Sjölin, Bo 1972-3. "Het Grote Oudfriese Woordenboek. Terugblik, balans, problematiek". In: *Us Wurk* 21-22:193-206.

Stallaert, Karel 1890-1977. *Glossarium van verouderde rechtstermen, kunstwoorden en andere uitdrukkingen uit Vlaamsche, Brabantsche en Limburgsche oorkonden*, vol. I-III (vol. III by Frans Debrabandere). Leiden-Handzame.

Vries, Oebele 1990a. "Zu einem Wörterbuch der altfriesischen Urkundensprache?". In: *Aspects of Old Frisian Philology*. Eds. Rolf H. Bremmer Jr, Geart van der Meer and Oebele Vries. Amsterdamer Beiträge zur älteren Germanistik 31/32; Estrikken 69. Amsterdam-Atlanta-Groningen (483-494).

Vries, Oebele 1990b. "Die Verdrängung der altfriesischen durch die niederländische Schreibsprache". In: *Niederdeutsches Wort* 30:83-96.

Vries, Oebele 1992. "Die Terminologie des schiedsgerichtlichen Verfahrens in der altfriesischen Urkundensprache". In: *Friesische Studien I. Beiträge des Föhrer Symposiums zur Friesischen Philologie vom 10.-11. Oktober 1991* (=NOWELE, Supplement vol. 8). Eds. Volkert F. Faltings, Alastair G.H. Walker and Ommo Wilts. Odense (29-52).

Vries, Oebele 1993. *'Naar ploeg en koestal vluchtte Uw taal'. De verdringing van het Fries als schrijftaal door het Nederlands (tot 1580)*. Fryske histoaryske rige 9. Ljouwert/Leeuwarden.

Vries, Oebele 1996. "*Also dat wy habbet ferdban kloppet*. Die Terminologie des Friedebannes bei der Eigentumsübergabe von Immobilien in der altfriesischen Urkundensprache". In: *A Frisian and Germanic Miscellany. Published in Honour of Nils Århammar on his Sixty-Fifth Birthday, 7 August 1996* (=NOWELE, vol. 28/29). Eds. Adeline Petersen and Hans F. Nielsen. Odense (251-271).

Wörterbuch der mittelhochdeutschen Urkundensprache. Auf der Grundlage des "Corpus der altdeutschen Originalurkunden bis zum Jahr 1300" unter Leitung von Bettina Kirschstein und Ursula

Schulze erarbeitet von Sibylle Ohly und Peter Schmitt. Berlin, 1986-...

Encyclopaedic Aspects of the Historical Dictionary and the Question of Context

Jeffrey L. Singman

Like many of the finer things in life, this paper is the product of a barrel. During my first year as an editor at the *Middle English Dictionary* (*MED*), I had to deal with a list of military gear which included a 'habergeoun barel'. I duly looked up the term in the *MED*, and found it glossed under **habergeoun** as '?a cylindrical framework on which a habergeon is set'. The more I contemplated this gloss, the less convincing it seemed. Habergeons are depicted in the Bayeux Tapestry suspended on rods, but I could see no obvious purpose for setting a habergeon on some sort of barrel – not to mention the obvious fact that barrels are not generally the right shape or size to slip inside a habergeon.

If the *MED* gloss for the habergeon barrel seems counterintuitive, the gloss for **harneis barel** [sv. **harneis** n. **6.** (a)] is not much better. This gloss, less than a hundred pages later in the same volume, describes a harness barrel as 'a barrel for washing harness'. Having already looked at **habergeoun barel**, I at first presumed that 'harness' here meant 'armour', and found the idea dubious. Washing implies water, and given the ease with which carbon steel rusts and the impossibility of drying a surface of tiny and densely inter-linked rings, one must suspect that washing would be an unsatisfactory means of maintaining one's armour. On examination of the *MED* files, I found that the editor in question had in fact been thinking of equestrian tack. The editor may reasonably be forgiven for overlooking the possibility that the harness in question was

armour, but the discrepancy between the **habergeoun barel** and the **harneis barel** surely ought to have been rectified in the process of review.

I mentioned this question to a friend of mine, a graduate student in Art History with some grounding in medieval material culture, on the off-chance that he might be able to suggest an alternative. He responded without hesitation that it sounded like the sort of barrel used to clean chain mail, and lent me a copy of Charles ffoulkes' *The Armourer and His Craft* for further information. I looked in the index under 'barrel', and found a reference with ease. According to ffoulkes, mail armour would be sealed in a barrel with sand and vinegar; the barrel was then rolled or shaken to clean the armour. ffoulkes offers a number of documentary examples, and even includes a late sixteenth-century painting in which a chain-mail shirt may be seen draped over the top of a barrel. Comparable devices are still used by museums for maintaining their armour collections.[1]

It would be unrealistic to expect so massive a project as the *MED* to be without its shortcomings, but we must at least ask whether these errors might reasonably have been avoided. Should the habergeon and harness barrels have been correctly identified? The question raises the fundamental issue of how far a lexicographer should be expected to delve into the cultural context of any given word, and how much of that context should be presented in the final gloss.

Obviously, this question relates principally to vocabulary that is in some way technical, or at least deals with specific cultural constructs peculiar to the society whose language is under examination. My examples here derive primarily from the field of material culture. This choice is based on two principal reasons. The first is that material vocabulary is relatively easy to explain, as opposed to that of an abstract technical field like law – a habergeon barrel is an easier concept to grasp than is a view of frankpledge. My other reason is that the vocabulary of material culture is typically one of the areas of a lexicon worst served by historical dictionaries. This focus on material culture might make the examples seem to some readers rather trivial – some may wonder whether this concern over the correct gloss for a 'habergeon barrel' is

[1] See Charles ffoulkes, *The Armourer and his Craft from the XIth to the XVIth Century* (New York: B. Blom, 1967). I am indebted to John Vernier for his assistance in this matter as in numerous similar questions I have encountered as a lexicographer.

something of a tempest in a teapot. To this I would offer two responses. Firstly, the current intellectual interest in the construction of identities must ultimately require us to examine more closely the most mundane and ordinary sorts of cultural phenomena, which play a profound role in shaping those identities. At the same time, it would be inappropriate for the editor of an historical dictionary to conclude *a priori* that a certain field of activity is so insignificant that one need not pursue appropriately professional means to gloss its lexicon as accurately as possible. And, as we shall see, the essence of my argument lies in appropriate professionalism, rather than in niggling precision.

In the lexicographic process, the research precedes the gloss, but from a structural perspective, the gloss takes precedence. The dictionary exists as a meaningful document at the point of interaction with the reader, and this interaction transpires at the level of the gloss. Consider the *MED* gloss for

saucefleume (a) A skin ailment considered symptomatic of a type of leprosy originating in a humoral disorder caused by sexual and dietary excess and characterized by red and black pustules swelling loss of hair, etc.

I would consider this an extremely good gloss for an item of technical vocabulary. The concept, which lacks a precise modern equivalent, is given a fairly full explanation, including typology, symptoms, and putative causes, yet the description remains within a glossal rather than a fully discursive format. It contrasts markedly with the *MED* gloss for

shirte (a) A garment for the upper body worn next to the skin by both men and women, a shirt.

Given that most readers of the *MED* are already familiar with the concept of shirts, this entry offers little in the way of useful information. 'An upper-body garment worn next to the skin' could also apply to a T-shirt, camisole, or tank-top. The entry implies that a medieval shirt was essentially equivalent to its modern counterpart, and may indeed be said to be misleading, in that the medieval shirt was not strictly an upper-body garment, as is a modern shirt: men's shirts covered at least the

crotch, and in some cases the thighs; women's shirts were essentially full-body garments, reaching from the neck to the lower leg.[2]

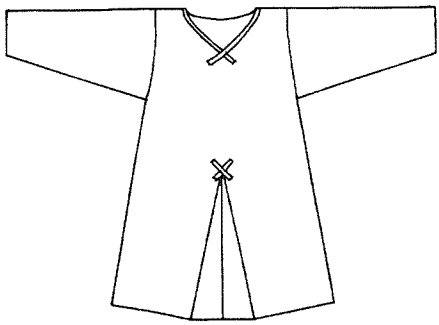
I would argue that an historical dictionary ought to be willing to offer its readers a bit more help, and to take a rather broader approach to its own function. The reader looking up a technical term in such a dictionary is probably not interested in the word merely as an element in a linguistic system, but wants some guidance as to the actual nature of the phenomenon to which it refers. We must after all bear in mind that our readers, unlike the readers of modern dictionaries, have no personal familiarity with the culture in question. At the same time, the reader does not necessarily have ready access to the relevant informational resources. We as lexicographers are extremely well placed to serve as channels of information, having the apparatus and the resources of our lexicographic projects behind us, and the reader may reasonably expect us to make full use of our position to provide more complete information than is often the case. Finally, we should remember that the historical dictionary is perhaps the single most powerful reference tool available for any historical culture, and as information professionals we should be committed to making that tool as effective as is reasonably feasible. In short, I would suggest that the historical dictionary ought to have a certain encyclopaedic dimension: as historical lexicographers, we should be prepared to include in our entries as much information as our evidence will sustain and as will fit within the lexicographic format.

To offer a concrete example, the entry for **shirte** could have offered the additional information that this garment was normally of linen or silk or a similar material. A more adventurous lexicographer might have added that the garment was pulled over the head, rather than buttoned in the front (in contrast with many other upper-body garments of the period, and with a modern shirt); one could conceivably even have mentioned that the seams were straight rather than tailored (as remained the case with nearly all shirts until quite recently). Another possible addition would be a reference to a useful secondary source (a technique occasionally applied in the *MED* already, albeit very rarely). In fact, I would go so far as to suggest that a superior historical dictionary would even include sample illustrations in some circumstances, if a useful contemporary example or illustration survives. Admittedly, a picture of a

[2] Jeffrey L. Singman and Will McLean, *Daily Life in Chaucer's England* (Westport CT and London: Greenwood Press, 1995), pp. 98-99.

thirteenth-century French shirt might be of questionable use in illuminating a fifteenth-century English text, but it would at least serve as a point of orientation, and would certainly constitute an improvement over any modern preconceptions regarding shirts which a reader might bring to the text. Such an entry might look like this:

shirte A garment worn by both men and women next to the skin, covering the arms and torso, and in some cases part or all of the legs, typically straight-cut and pulled over the head, and made of linen, silk, or a similar washable fabric, a shirt.



A Shirt (French, thirteenth century).

This kind of entry goes significantly beyond the traditional content of a historical dictionary for words of this sort, but we should also bear in mind that the traditional content is predicated on a paper dictionary. No genre of text stands to gain more from the new technologies than does the reference work: SGML and the related text-marking languages offer the possibility of a multi-layered dictionary which can display information at different levels to accommodate both philological and encyclopaedic approaches to the dictionary.

Having suggested this rather radical vision of the role of the historical dictionary, I must return to the question of research, for clearly the precision of our glosses cannot be separated from the depth of the research on which they are grounded.

No amount of research can guarantee perfect accuracy. Even if perfection were not made impossible by the medium of language itself, we would still be faced with misleading sources, incomplete information, and a host of other obstacles. What we must look for instead is the kind of accuracy which arises naturally from appropriate lexicographic practice. This leaves us with the delicate task of drawing the line between appropriate and inappropriate.

Ultimately, the appropriateness of any intellectual practice must be defined in relation to the pre-existing intellectual resources. A theory is clearly deficient if it fails to take into account established ambient data; it is equally faulty if it ventures past the ability of the intellectual milieu to sustain it – regardless of whether it is in fact more accurate. Similarly, an appropriate lexicographic practice is not one which always produces accurate results, but one which makes suitable use of the available resources.

The principle is well illustrated by the examples of the habergeon barrel and the shirt. The question of the habergeon barrel would have been easily resolved by consulting a knowledgeable specialist – it did not even take a significant amount of my informant's time or effort, as he could cite the relevant information off the top of his head. Alternatively, this information could have been acquired by consulting ffoulkes' book on armouring – by no means an obscure text in its field, since it has been around since 1912, and is now in perpetual reprint with Dover Books. It is no coincidence that this lexicographic failure occurred in one of the areas where the *MED*'s reference collection is most shamefully weak: we have a grand total of four works on arms and armour, none of them more recent than 1934, and only one of them specifically on the medieval period. The weakness of our entry for **shirt** might likewise be connected with the paucity of our costume sources. This collection is more substantial, numbering some 17 volumes, but only about half of these are of any use: the rest either represent nineteenth- or even eighteenth-century levels of scholarship, or are superficial treatments of costume geared for theatrical use; some deal only with post-medieval clothes. The very significant Scandinavian bog-finds of medieval clothing and the rich archaeological studies published in recent years by the Museum of London are wholly absent. By contrast, the admirable quality of the entry for **saucefleume** may have something to do with the 46 volumes on medicine and medical history in our reference collection.

I would suggest, then, that one of the first stages of any sound lexicographic practice would be the systematic identification of a body of resources, both human and textual, upon which to draw, across the range of anticipated semantic fields. Let us break this process up into its component parts:

1) Identifying the range of fields. This task would be challenging, but would be simplified by reference to the Library of Congress system

for cataloguing books, which offers a ready-made framework for the classification of human knowledge from which the dictionary could select appropriate headings.

2) Identifying human resources. This depends on the willingness of other scholars to assist the lexicographers in their task, but I doubt that this would prove much of a problem. Scholars in the humanities are as a rule only too happy to share their expertise and perspectives. After all, I am not proposing that these consultants should actually engage in research on the lexicographers' behalf – rather, they would be asked to offer insights based on their pre-existing knowledge (as did my art historian friend in the case of the habergeon barrel). Given the sustained torrent of this kind of activity already taking place on the Internet, I think there would be no difficulty in finding plentiful assistance.

3) Identifying textual resources. This would, in all likelihood, be the most demanding of the tasks I have proposed, yet I believe it would prove less difficult than might at first appear. For a start, a helpful and easily accessible bibliographic base already exists in the bibliographies included in the Dictionary of the Middle Ages (although these bibliographies will become less useful over time). Secondly, if one has already found a consultant for the field of activity in question, that person will probably be able to recommend a basic bibliography with minimal effort.

Rather more difficult would be the question of how far to extend this bibliography under any given heading. I would suggest that the inclusion or exclusion of any given book would depend on two principal factors. The first is the informational richness of the text itself. A text well-executed and well suited to the needs of the lexicographer comes with an *a priori* reason for inclusion. The second is the anticipated needs of the field of activity in question. An extensive field with a large number of words would require a larger number of sources, and a field in which there are fewer textual resources available would be more inclined to need any particular text. Again, to offer a concrete example, I believe that the size of the *MED*'s costume collection, and perhaps even its armour collection, would have proven quite adequate, had the texts been better chosen. The decision would ultimately rely on personal judgement, but in fact these sorts of decisions are already being made when a dictionary assembles its reference collection, or when an individual lexicographer decides how far to research any particular

word. I am merely suggesting that the decision be made more reasoned and explicit.

The existence of these sorts of resources would put the efforts of any individual lexicographer on a much surer footing. Historical dictionaries are notoriously the work of philologists. When it comes to etymology, phonology, or semantics, we are in our element, but faced with a point of law, a question of theology, or a common household implement, we are at a loss. A sound structure for the pursuit of lexicography would be one in which the lexicographers had a well-defined and well-chosen corpus of resources for dealing with words outside of their fields of personal expertise. The lexicographer would be expected to pursue any technical vocabulary in the textual corpus offered; if this failed, the next step would be to turn to the consultant. I would not wholly circumscribe the lexicographer's action – there is no reason why one might not pursue the word even further on a hunch or even on a whim; but the principal point is that lexicographers who had consulted the official resources would have discharged their duty as responsible lexicographers. If these sources are chosen intelligently to include the best and most suitable of available resources, we can be reasonably confident that the glosses produced will represent a satisfactory level of lexicographic scholarship.

Having said these things, I must temper them with a certain melancholy qualification. One may wonder whether these suggestions do not come too late. For the *MED*, the time to have undertaken this kind of structural work was half a century ago, not as the project enters its final years. Not only is the *MED* nearing completion, but the entire field of historical lexicography seems to be in its twilight: funding is scarce in general, and the current intellectual climate seems at best indifferent, at times openly hostile to the aspirations of philology. To offer a blueprint for the initial stages of a lexicographic project might seem rather futile.

However, I would like to think that the long-term prospects are not quite as bleak as they may appear. To begin with, we are in the midst of an information revolution whose full impact on academia remains to be felt. The accelerating advances in computer technology create a strong incentive for the pursuit of activities which will restructure information in such ways as to make large volumes of data more readily usable to the individual researcher. Academia will only really have entered the computer age when it begins to allocate significant resources in this

direction, and the dictionary is a prime example of an information-restructuring resource.

Even if the plan for dictionary work I have here outlined will never have the chance to be implemented, I believe at least that the principles I have enunciated retain some relevance for individual lexicographic practice. We must always remember that an historical dictionary is not a timeless artefact whose purpose is to encapsulate in perfect suspension a language of the past. Rather, it is a living tool, and must be judged in relational terms. Its quality must be measured in part against the quality of the pre-existing tools available to create it. For a dictionary, to err is excusable, but to fail to make use of the best available resources is not. But above all, the dictionary will be judged in relation to its future audience. In composing an entry for any word, the needs of the user must always be foremost in our minds, and in the case of technical vocabulary, this means offering enough information in our glosses to provide the reader with some orientation in what is, after all, an alien culture.

Identification of the Veins and Arteries in Middle English Medical Literature

James M. Girsch

In deciding upon a topic for this paper on historical dictionaries, I considered the first requirement to be the presentation of a problem that would have aspects of interest both to the professional lexicographer and to medievalists from other disciplines – that is, a problem presented by a word or group of words whose treatment forced me to undertake some exploration outside the usual boundaries of my work and to enlist the assistance of others. The second requirement, it seemed, was to present some rather extensive lexicographical work that had necessarily remained unfinished but that would stimulate further research by demonstrating some interesting difficulties, ambiguities, and of course, errors, for comment and corrections, which such work by its nature must always bring with it.

Lexicographical treatment of the Middle English noun *veine* came to mind instantly as a good example for a discussion of this sort. Though it presents few problems in syntactic or semantic behaviour, morphology or etymology, its prominence in medieval medical literature – a subject in which interest seems to be growing – the significance of its occurrences for our understanding of the history of physiology and surgery, its presence in a species of text whose study may cast much light on the history of translation as well as that of its own particular field, and finally the baffling ambiguities and other difficulties of some of its collocations make it a particularly attractive candidate. Furthermore, its obvious importance in the modern world of anatomy and medicine means that it has been thoroughly examined and treated lexicographically in modern medical works, so there should be a good

number of modern equivalents for the medieval terms, the problem being in many instances simply to *find* the modern equivalents and to assign them.

So I will present here a work in progress: the *Middle English Dictionary*'s treatment of a rather commonplace noun in one of its special sub-senses, restricted to its anatomical applications. The material collected and glossed in this article will I trust be useful in itself, in spite of its provisional nature. What I hope to accomplish as well is to provide a glimpse into the workings of our historical dictionary project, treating some of the problems that we encounter and of our ways of finding solutions, as well as making clear some of the limitations inherent in the work, which in turn do suggest the proverbial opportunities for further research. In this instance those opportunities are fascinating and abundant indeed.

I have placed at the end of this article a treatment of the word as it now stands, edited by me for the *Middle English Dictionary* over the course of many weeks and now sitting in its cabinet at the University of Michigan awaiting review-editing. The material is set out in a form very close to that which it will have in the finished *MED*, and readers are advised to have a copy of the *MED* and its *Plan and Bibliography* at hand if they plan to make use of the material. Moving from the text of the article to the collocations and glosses is indeed somewhat unwieldy, but I expect that the reader will be able to refer to glosses easily enough while ploughing through the article if they are placed at the end; a photocopy of these final pages in hand while reading will be a useful aid.

I want especially to address a set of problems that I think are characteristic of this kind of technical term. I will be concentrating on the anatomical senses of *veine* – labelled Anat. in the *MED* style – which comprise all the glosses under Arabic numeral 1, but I have printed the entire semantic treatment of *veine* in the interest of completeness (Appendix One). The accompanying illustration showing both the systemic and the pulmonary circulatory systems will clarify some of the problems under discussion (Figure One).

Please bear in mind that in the printed *MED* all the senses and sub-senses here will be equipped with quotations to support the glosses. There is also a good possibility that new material will come to light and that the reviewer will discover some discrepancy or conflict or even some new collocations to be added to what I have here. Thus it is

important to recall that the glosses presented here are not in their final form, and that the final version is not likely to be available within a year or two. It may well be an improvement and expansion of the word as it now stands. In the meantime, however, there is no reason not to accept what I have presented here and to make use of it if there is need. It was done with unusual care and attention to detail, as I knew that I was working in a field of study that ordinarily requires much more training than I could lay claim to. My colleague and friend, Dr Robert Mory, an *MED* editor with much more experience in the field, did me the great favour of reading carefully the finished work; his *imprimatur* gives me enough confidence to recommend what is here, at least provisionally.

The general setting of what I am about to discuss is provided by the *Middle English Dictionary*'s particular way of assigning material to editors: each editor is given a box-full of material (the boxes are of a standard size) representing a list of words in alphabetical order across a short range of the alphabet – *variaunt* to *velti-fronde*, for example; or a box containing only *truth* (how well I realise what such a box might be worth in a better world) a truly formidable word for the lexicographer to essay, plus a tiny fistful of other words. Thus there is tremendous variety in the content of boxes, and the work that any single box represents is very far from being consistent. Some boxes take two months; another might take over a year. There is as well some direction of certain kinds of words to editors with particular interests or training. We have on our staff, for example, editors with extensive formal training in Old English philology and semantics, some with Old French, some with much expertise in Medieval Latin – the Latinist might be handed a box that abounds in Middle English words from Latin and Old French and the Anglo-Saxonist will probably be given the big reflexes of Anglo-Saxon words if possible – the dreadful and titanic modal *willen* comes instantly to mind – but otherwise any editor can expect to face just about anything at any time. I do not know how other historical dictionaries divide up and allocate the material, so I am unable compare our methods, but that is beside the point. What happens is that our editors, like readers and students in general, must come to grips often with material that is unfamiliar and in a severely limited period produce a treatment of the word that is comprehensive yet concise and as accurate as possible. Perhaps in some instances an editor's ignorance of the details of a word's context is a positive virtue – perhaps it allows the editor to

approach the task with the same questions that non-specialist readers will ask.

One aspect of the *MED*'s procedure is of primary importance for both the main issues that I want to treat here: when an editor is assigned a box of raw material, the *MED*'s presumption is that no further *systematic* collection of additional material will be undertaken. That is, what arrives on an editor's desk is, with a few minor exceptions, all that an editor is expected to use in the treatment of a word, though we have some very limited time and freedom to search for more quotations if we feel that there is good reason to do so. This setting of limits is essential for a project of the scope of the *MED*, necessarily working under a deadline – but it gives rise to problems, which I will discuss below.

The most pressing questions that an editor faces in treating a word like *veine*, then, are 1) how can an editor untrained in anatomy and medicine and possessing only the rudiments of knowledge of the history of medicine treat properly and accurately the technical terms of those fields of knowledge? And 2) to what extent, considering the relatively small amount of time we can afford to spend on any word, should a generalist attempt to be on the one hand comprehensive and, on the other, specific in treating the obscure or difficult occurrences? Bear in mind here that the basic semantic distinctions were fairly simple to establish: we are really only concerned with sense 1, and the basic distinctions are (a) a blood vessel, a vein or an artery; (b) a vein and (c) an artery. The real difficulties arise in the handling of the numerous collocations in sense 1. (b) To rephrase the question, treatment of a word like *veine* seems to require an approach more encyclopaedic than is our usual approach given the constraints placed on us by the usual forces, necessarily so when treating a term from medicine, law, military affairs, and similar technical lexicons. Where does one draw the line in doing outside research? How much ambiguity is tolerable in the preparation of a work that must serve as the basic research tool for all scholars working in Middle English for years to come ?

In spite of the word's very obvious significance in its medieval context – given the prominence of bloodletting as both a therapeutic and a prophylactic procedure – the present state of scholarship in Middle English medical texts leaves the lexicographer with very little to turn to from experts, for example editors of medical texts and historians of medicine. Despite the appearance of such invaluable work as Wallner's on-going edition of the New York Academy of Medicine MS 12

(*Chauliac (1)*, as we call it), which is equipped with a glossary that has been helpful at times, there is no detailed treatment of the English nomenclature of the veins and arteries now available.[1] It seems likely that the materials of the *Middle English Dictionary* contain the best and fullest collection of the important source-texts available anywhere at present. For the draft of *veine* (again, considering only the anatomical sense and ignoring senses 2 and 3), the *MED* supplied quotes from about 50 texts, including some texts available only in manuscript or dissertation form. I had in addition complete access to the unused materials from all previous volumes. So the best opportunity likely to present itself for a description of *veine* as systematic and complete as possible was almost certain to occur during the *MED*'s handling of the word. It seemed, in other words, that I was left with the task, or opportunity, of producing the most comprehensive lexicographical treatment of the difficult anatomical occurrences.

Furthermore, in spite of the claim to some kind of completeness of sources that I have just advanced, I can at the same time point to the abundance of material that our reading program has of necessity overlooked. I have in preparation a list of all works in the *MED* Bibliography that might contribute to a more complete picture of *veine* – the total is somewhere near 85 works. I have no clear picture of the varying degrees of thoroughness with which this collection of sources was searched for occurrences of *veine*, before the material came to me, but the additional material that I found in a single work, which I will discuss in a moment, should provide some idea of the limitations of a large non-specialised dictionary such as the *MED* in the handling of specialised vocabulary. Finally, there is undoubtedly a wealth of unedited material that would both resolve desperately opaque collocations and add, one must assume, gratifyingly large numbers of new ones. Thus any handbook likely to be produced within the next few

[1] Thus far Wallner has produced edited versions of Books I to V of the *Great Surgery*, as he has titled it. The titles and essential publishing data vary with each Book, and it is so complex that to repeat it here would be tedious in the extreme. Readers will find ample information on the various volumes in the *MED's Plan and Bibiography* (University of Michigan Press, 1954) and its *Supplement I* (1984). The single volume published after the Supplement was printed (1988) contains Book I and will easily be found with the other volumes in any library.

years will, it is to be hoped, be superseded by another or by a supplement as new works are edited and studied.[2]

Once I had gathered all the occurrences of *veine* and finished my glossing (indeed at a point some two years after I had gone on to other problems) I returned to *veine*, and as a simple test of the *MED*'s comprehensiveness, I chose an important, easily accessible EETS edition – one of course lacking in a glossary, index or notes – and read it systematically over a period of two weeks, looking for occurrences of *veine* that our reading program had missed. Here it should be noted that the unique Paris MS translation of Chauliac's *Cyrurgie* was read by the *MED* for its contents before Margaret Ogden's excellent EETS edition appeared in 1971.[3] That is, a group of readers ploughed through the text looking for important words to be added to our boxes, but they did so from a photostat of the manuscript. This work is a chore sufficiently grinding when one has a nice neat printed book to stare at, but it must have been heavy going for those who harvested occurrences of the word *veine,* not to mention the other words of every variety. A second volume with notes and glossary, based on Mrs Ogden's notes, is now on the verge of completion by Prof. Michael McVaugh, whose name will be familiar to anyone who has read even a little of the history of medicine, and Dr Mary Jane Williams, a review editor at the *MED*, its able and prodigiously knowledgeable bibliographer, and a former associate of Margaret Ogden. The appearance of this volume will be very welcome to any student of the history of English medicine, and especially so to

[2] Prof. Linda Ehrsam Voigts, one of the most prominent experts in the field, in a private communication to me (March, 1995) informs me that there are at least two very important texts containing material for a comprehensive treatment of circulatory anatomy in Middle English, both of which must be taken into account before even a provisional work that seeks to achieve some completeness is produced. The first has been edited as a thesis by Maria Jose Carrillo-Linares from Trinity College Cambridge MS R.14.53, ff. 40v-44; the second is the text of Bernard of Gordon's *De prognosticiis*. I had no access to nor knowledge of these materials when this article was in preparation, and any further work that I undertake will certainly survey both works and incorporate whatever pertinent information they contain. Here I would like to offer my thanks to Prof. Voigts for provision of these references as well as other helpful information that she has shared with me.

[3] *The Cyrurgie of Guy de Chauliac*, ed. Margaret Ogden, EETS OS 265 (New York, 1971).

those who have tried to use this dense endless work with no index or glossary to guide them – like me.

In order to give some idea of the scope of the material that the *MED* necessarily overlooked in its treatment of *veine,* I will present here a selection of some of the more interesting occurrences of *veine* and related words; that is, words for specific veins and arteries that I have found, missed by the *MED*'s reading program, all of which will be incorporated into the *MED*. Remember that the word has not been reviewed, so there is plenty of time for these citations to be added. I have omitted some for the sake of brevity, of course, and others here (like *subetales* or *apoplecite)* that are glossed either elsewhere in the *MED* or in other editions, especially Wallner's edition of the New York Manuscript. Nevertheless, these significant occurrences, it must be remembered, were utterly lost to me until I went to look for them *after* I had finished the formal *MED* treatment of *veine* – that is simply what our method and limited amounts of time and money require. References are to page and line of Ogden's EETS edition of Chauliac's *Cyrurgie (Chauliac (2)).* Amongst the most intriguing collocations:

1. 47/19 "deep guidedge, appoplectices, subetales": note that *apoplecite, apoplice* are glossed by Wallner in *Chauliac (1)* as "carotid arteries"; *guidegy* is glossed there as "jugular veins". The entry *subetale*s is glossed by *MED* as "Anat. the carotid arteries".
2. 68/8 *emoroydales* as noun: "there ben 5 braunches of veynes applied (around the fundament) cleped *emoroydales*".
3. 68/27 *streyning veines* and *berynge arteries*: I do not yet know what these might be.
4. 134/35 "cut the *temple veins* and the *forhed veins*": I do not have "temple veins" in my treatment as it now stands. Note that *veine of the forhed* is glossed in *MED* as "one of the two supratrochlear veins" – is this the same thing?
5. 171/33 "*schiracioun* is cleped *swelling veins* without kind in the testicles": this is glossed in *MED* as a *hapax*, from this locus: "a hardening of the flesh in or near the testicles".
6. 257/12 "of the cutting of the *spermy vein* – it is but a jape or a trufle": could this be the spermatic duct?
7. 257/24 *heads of the veins* and 322/15 "hemorrhoids are swellings with humors in the heads of the veins": note the

rather confused physiology concerning the stoppage in the "heads of the veins".

8. 385/26-32 "For lepre, blode last of the great veins ... and opening of the small veins of the face may accord": there is no specific collocation, but further on, "Haly Abbas commands the blode last of the first two *great veins* and the *veins behind the ears* ... and of the *mediastines*; Rasis begins at the *vena purpurea* of the right arm". *MED* had this one sense of *mediastine* n., though the time I spent in indirectly finding it suggests the sort of work that will make a comprehensive treatment quite a lengthy undertaking. See *MED* at *mediastine* n. "(a) Anat. the mediastinum (the septum that divides the thorax); (b) ? the antecubital vein of the arm".
9. 417 "for skalles: let blood of the vena cephalica and the *veins behind the ears*".
10. 435/22 "gutta rosacea cure with bloodletting of the *veins of the forhed and the nose*": should these be glossed, for example?
11. 440/28 "Galien grants cutting of arteries of the temples for inner tears and *veins of the temples* for outer tears": this remains opaque to me.
12. 454 "*Sebel* after Avicenna is a pannicle coming in the eye of the swelling of veins thereof in the outermost of the conjunctyf and cornea – some comes by the *inner veins* and some by the *inner uttre veines*": this is obviously very confused; there are several occurrences here, but no really specific collocations. But cf. *sebel* n. in *MED*: "an eye disease characterized by veins engorged with blood and with a web obscuring the sight".
13. 485/10 on cure of the tooth-ache: "bleed *cephalica* and *veins of the lips and tongue*": these veins of the lips and tongue could simply be any localised vessels that the practitioner finds easy of access, or they may be specific anatomical entities.
14. 535/35 the *forthright vayne in the forehead* (for headache in back of head): again, this defeated any attempt to assign a modern equivalent to it.
15. 538/5 "The fourth thing is by which *veines* that blode last is to be made". Haly Abbas (in his Real Book) gives 33:

> 12 in the arms: 2 *medynes* [vr. of mediana n.] 2 cleped *cephalice*, 2 *base veins*, 2 under arm holes; 2

under elbows and the 2 *seynales* [these latter are the *seinales* n. pl., glossed in *MED* as "Anat. Veins on the back of the hand."]

13 in the head: 2 behind the ears, 2 in the corners of the eyes, 2 *hollow veins*, 2 of the *hattrelle*; the [single, ut vid.] *vein of the forehead*, the *vein of the li3te*, the *vein of the nose*, 2 *veins under the tongue*.

8 in the feet: 2 in knees, 2 in the ankles, the 2 *splenes* and the 2 in the wrist of the foot.

16. 538/25 "veins for the blode last of the elbow are 3: *within, without* and the *middle vein*". The vein "within", the text suggests, helps parts from the neck downwards; that "without" helps in the places above; the "middle" helps either. These are difficult to assign to modern equivalents.
17. 540 "*common veins* cut by length; *particular veins* cut overthwart; arteries be hyndeng and by cautery": this obviously will bear some closer examination.
18. As an additional example, compare Jasin's edition of the *Liber uricrisiarum*, p. 155, where the Latin *vene juveniles* is translated by "*3onge veyns* or ellys *fryk veyns*".[4]

How, then, can an editor with no medical training handle with any confidence such terms from the field of medicine? The problems I encountered in producing a coherent picture of the Middle English word *veine* bring into focus areas of knowledge that are in need of supplement not only for the specialist in Middle English, but also, and especially, for the student or non-specialist who is trying to make sense of a particular text or body of literature. Undoubtedly, any attempt to provide an accurate guide that assigns the Middle English named veins to their modern English counterparts will have to depend on the assistance of a well-trained expert in human anatomy who has an interest in history.

A good example of the vexing problems to be solved in this regard comes from the phrase *arterial veine*. As I have glossed it, the collocation appears in two sub-senses, and even in the first of the two there is some ambiguity, resolution of which depends upon the Middle

[4] *A Critical Edition of the Middle English Liber uricrisiarum in Wellcome MS 225*, ed. Joan Jasin. Ph.D. dissertation (Tulane University, 1983).

English text that one is reading. *Arterial veine* seems to denote any of the following:

1. One of the four pulmonary veins, or their juncture at the Left Atrium.
2. The inferior Vena Cava.
3. The pulmonary artery or the Left and Right branches of the pulmonary artery.
4. Yet another possibility raised by The Paris MS of Chauliac's *Chirurgie,* or *Chauliac (2)* in *MED:* the aorta, which sense is not provided for in the glosses as they now stand.

The adjective *arterial*, I should add, is glossed in *MED* as follows verbatim: "Having the characteristics of an artery – applied to the pulmonary artery and to the pulmonary vein". Only two quotations are given, both from translations of Chauliac's *Surgery,* one in support of each sense.

The following passage from Ogden's EETS edition of the Paris text of Chauliac's *Surgery* (*Chauliac (2)*) – which I found only after I had finished a draft of *veine* n. – demonstrates the likelihood that further research will uncover more complexity than is apparent now. Here I am quoting from Wallner's edition of *Chauliac (2)*, 55/11 – 56/7:

> And therfore there beeth in hym (the heart) two mouthes. By the ri3te mowth gooth yn the braunche of *ascendynge veyne* beryng vp the blood fro the lyuere [this is the inferior vena cava] and it gooth oute at the same. Oo parte is cleped *vena arterialis*, the *wosy veyne* [this must be the *pulmonary artery*] the whiche gooth to norisshe the longes .... And *vena pulsatilis,* the *pulse veyne*, goth oute fro the lefte mowthe [this is the aorta]. Wherof oo party goth to the lunges, the whiche is cleped *venalis*, beryng stynkyng vapoures to the lunges [this appears to be the juncture of the pulmonary veins] .... Vppon the self herte the lunges moven to be coled. Whos substaunce is softe ... spongyouse and white. Withynne the which thre kyndes of vesselles ben plaunted togedre, that is to say the braunche of the *veyne arterial*, the which veyne springeth (as it is

> said) of the ri3t ventricle of the herte [this is the pulmonary artery] and the braunche of the *arterial veyne*, the which braunche cometh of the lefte ventricle and with ham beth braunches of the *trache arterie* (i. the *thropul*) beryng to hym ayre for the herte. The whiche vesselles forsothe beth departed by all the lunges vnto the lefte parties.

I would like to note here two particularly bewildering ambiguities and conflicting uses that, again, appeared only after I had finished the glossing based on the materials provided.

1. *arterial veyne* is very difficult here:

A: on the one hand it would seem to be the juncture of the pulmonary veins, as I have glossed it in 1. (b) since here they are said to arise from the left side of the heart and to be connected to the lungs – though they enter at the left atrium rather than the ventricle in which case *ventricle* is either loosely or erroneously used;

B: on the other hand, it could be the aorta, since it does spring from the left ventricle and does pulsate (as an artery) – though the aorta is *not* connected with the lungs.

2. *pouse veine*: in this passage it seems to mean the aorta 1. (c); but it is also glossed under 1. (b) as "one of the four pulmonary veins or their juncture at the left atrium of the heart".

Other ambiguities that remain in the *MED*'s treatment of the word point to these general areas in need of attention by scholars:

1. Lack of edited Middle English material with anatomical materials.
2. Lack of detailed glossaries for those that do exist.
3. Difficulty of finding exact and unambiguous correspondences between Middle English and modern anatomical nomenclature.

A good number of other questions of course remain, and these probably represent just the beginning of a number of interesting issues that the study of medieval English medicine and surgery will raise and eventually answer:

1. How "learned" was vernacular medical writing? (see Voigts *Middle English Prose*, 318).[5]
2. How "independent" of the Latin lexical system was English, which in most cases was translating Latin? (Voigts, *MEP*, 316) Is there enough independence to be significant (this is a corollary to point 1)? In the texts I have dealt with, how often is a Latin equivalent given? Is there evidence (as suggested in Jasin's article on *Liber uricrisiarum* in the *JEGP*)[6] that the way *veine* is handled suggests broadening of access of these works to reach the practitioners without university training? Of course, not only physicians let blood. Note the distinction in Middle English between *leche* and *physician*: Jasin notes this distinction in her *JEGP* article on the audience for the *Liber uricrisiarum*.
3. Does the finished word reveal anything new about the medieval understanding of what we call the circulatory system? Consider here the confusion between *venous artery* and *arterial veine*, which as you can see denotes the pulmonary vein (1.b), the inferior vena cava (1.b) and the pulmonary artery (1.c). It may also denote the aorta in (1.c). Where in Middle English texts is this discussed most fully? Knowing that in itself would prove useful.
4. Does the range of texts across the period show clearly any development of knowledge or conception? Or is the body of knowledge static within the Middle English period – is it still completely derivative? Is there any sense that anatomical knowledge is increasing or developing during the period? Answers to these questions would be most interesting.
5. What are the most authoritative texts? Is there one in particular that presents the most comprehensive treatment? Does the finished word bring any relationships amongst Middle English texts?

[5] Linda Ehrsam Voigts, "Medical Prose", in *Middle English Prose: A Critical Guide to Major Authors and Genres*, ed. A.S.G. Edwards (New Brunswick, 1984), pp. 315-35.

[6] Joanne Jasin, "The Compiler's Awareness of Audience in Medieval Medical Prose: The Example of Wellcome MS 225", *Journal of English and Germanic Philology* (1993), 509-22.

6. How many collocations did our treatment miss? Of course, this can only be estimated, but even an estimation would be useful.
7. It is evident from the list of works quoted for treatment of the anatomical sense of *veine* that a variety of works contributed to the full picture of medieval circulatory theory and practice, texts beyond the realm of medicine or surgery: the various *Secreta Secretorum* editions,[7] glossaries, the *Treatise on Horses*[8], Trevisa's translation of Bartholomeus Anglicus's *De proprietatibus*[9], and other similar works. What might a comparison of veterinary with human physiological theory tell us? Is there enough material upon which to draw? Would the differences be truly significant?
8. Where do medieval authorities differ? Are there two or more distinct "circulatory traditions" – say an English versus a Continental tradition? This would be very interesting. Does the material gathered under *veine* make anything like that clear?

Some solution of these problems and resolution of the various difficulties that I have described may be possible in the near future. If there is general agreement that a handbook of circulatory terminology would indeed be worthwhile, I hope to take what I have gathered and learned from my handling of the word and expand it to include the arterial system (which is smaller both in the human body and in Middle English medical literature) and to publish the results in an independent monograph. A comprehensive treatment of *veine* and *arterie* might provide some answers, whole or partial, and solutions to the following questions, many of which have been raised by leading scholars in the field such as Linda Voigts, Michael McVaugh, and others in recent years. I will close here with the suggestion that answers to those questions, which further research must address, may well increase our knowledge of not only the history of anatomy and surgery, but of the

[7] *Secreta secretorum: Nine Middle English Versions. Vol. 1: Text*, EETS OS 276 (Oxford, 1977). See also Robert Steele, ed., *Three Prose Versions of the Secreta Secretorum. Part 1: Text and Glossary*, Glossary by T. Henderson, EETS ES 74 (London, 1898).

[8] Ann Charlotte Svinhufvud, ed., *A Late Middle English Treatise on Horses, Edited from British Library MS Sloane 2584, ff. 102-117b* (Stockholm, 1978).

[9] *On the Properties of Things: John of Trevisa's Translation of Bartholomeus Anglicus's De Proprietatibus Rerum*, ed. M.C. Seymour et al., 2 vols (Oxford, 1971).

history of translations, textual transmission and that of the English language itself.

Appendix One

vein(e n. Also wein(e, vena, vene, vain(e, wain(e, wane; pl. veines, veinus, etc.
& veinnis, (early SW) veinene, & (?errors) veanx, vyne.

1. Anat. (a) A blood vessel; also in fig. context; also, fig.: the blood, the innermost seat of one's feelings, the heart; pl. the body; also, fig.: the essential meaning (of a statement) [quot. (1382)]; bresten the veines, to kill (sb.); everi – , alle the veines, every part, every bit; in everi – , in every part of the body; totally, fully; keenly, severely; sineu and – ;

(b) a vein as distinguished from an artery by location, structure, behavior or function; also, ? a venule or an arteriole [quot. ?c1475 1st]; – arterial, arterial – , one of the four pulmonary veins, or their juncture at the left atrium of the heart; also, the inferior vena cava [quot. c1475(1392) *MS Wel.564 40a/b; see also arterial adj.]]; – on the bak, a vein of the arm, suitable for bloodletting; – basilica, basilica – , the basilic vein; also, one of the four dorsal digital veins of the foot [quot. *MS Htrn.95 142b/b; see also basilica n. and adj.]; – of the elboue, ? the anterior or posterior ulnar vein; ? the median cubital vein; – epatica, the basilic vein [see also epatica n.(1)]; – eralis, ? = – erilis, ? the basilic vein; – of the forhed, ? one of the two supratrochlear veins; – of the gret to, one of the dorsal digital veins of the foot; – of the longes, ? the median cubital vein; ? the accessory cephalic vein; – meden, a condition producing a painful, swollen, tortuous vein in the leg; also, such a vein; – organik, a jugular vein [see also organik adj. (c)]; – porta, the portal vein [see also porta n.]; – salvatel (– of the splene), a vein between the ring finger and the little finger of the right hand [see also salvatel n. & splenatica n.]; – sophene, the small saphenous vein; also, ? the great saphenous vein [quot. a1500 1st; see also sophena n. & sophene adj.]; – of the wombe, a superficial vein of the foot; ? the medial marginal vein; veines capillars, capillar veines, ? venules; ? arterioles [see also capillar adj. (a)]; veines miseraices, miseraice veines, tributaries of the portal vein, including the mesenteric veins, the gastric veins and the splenic vein, and their tributaries [see also miseraice adj. (a)]; bas –, the basilic vein; capital – (– of the thoume), a dorsal digital vein in the thumb, an extension of the cephalic vein; pl. veins associated with the head; commune –, ? a venous trunk before its division into tributaries; coral –, – coral, the median

cubital vein; also, ? the cephalic vein [quot. a1400 Lanfranc 131/12]; eie –, an extension of the cephalic vein that joins the dorsal venous network of the hand; hed (cephalica) –, – (cephalica) of the hed, the cephalic vein [see also hed n. (1) 1c.(d)]; herte –, – of the herte, ? the median cubital vein; ? the median vein of the forearm; holwe (ascending, kile, ramouse) –, – ascending (concava, kile, ramosa), the inferior vena cava [see also holwe adj. 3.(c); kile n. (1) (a); ramouse adj.]; livere –, – of (on) the livere, the basilic vein [see also livere n. (2)]; maister –, ? a large or main vein; also, one of the saphenous veins [quot. a1500 1st; see also maister n. 1.(g)]; mediana (bodelich, middel) –, the median cubital vein [see also mediana, n. & middel adj. 1.(d)]; nekke –, ? an equine jugular vein; particuler –, ? a vein branching off from a larger venous trunk; pouse –, one of the four pulmonary veins, or their juncture at the left atrium of the heart; purpure –, – purpurea, ? a superficial antecubital vein suitable for bloodletting; ? the median antecubital vein [see also purpurea adj. as n.]; quik –, a living vein; menstrual veines veines menstrues, ? veins presumed to connect the uterus and the mammary glands; ? the two uterine veins and their tributaries; milken (milki) veines, ? veins presumed to connect the uterus with the mammary glands; ? chyle vessels, lacteals; ramouse veines, the hepatic veins; also in cmpds. and cmbs.: – blod, blood drawn in bloodletting; also, bloodletting [quot. (c1385)].

(c) an artery; arterie –; felen (gropen, tasten) – (sineu & –), to check the pulse of (sb.); also, ? to palpate (an organ, a part of the body) in diagnosis; – aborchi, the aorta [see also aborchi n.]; – arterial, arterial –, the pulmonary artery; pl. the left and right pulmonary arteries [see also arterial adj.]; – pulsatile, the aorta [see also pulsatif adj. (a)]; pouse –, – of the pouse, an artery; also, the aorta; wosi –, the pulmonary artery;

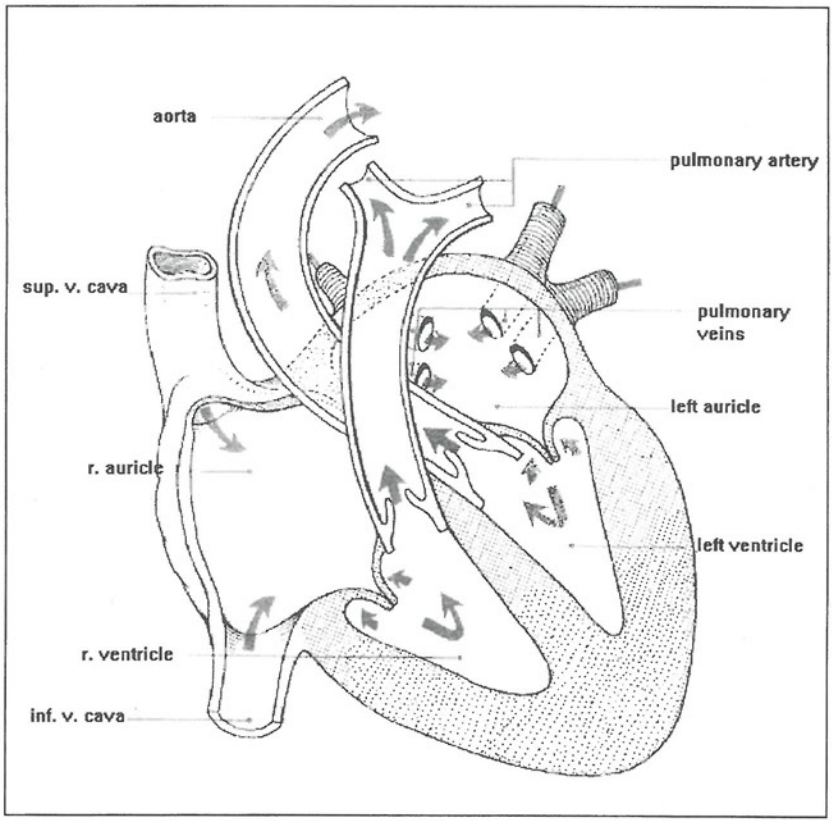
(d) a bodily passageway, meatus or duct; – kile, the ureter; veines of the urine, ? ureters; ? urethras; vocal –, ? the larynx; ? the fissure between the vocal folds in the larynx, the rima glotidis;

(e) a sinew, ligament or filament; pl. ? the vocal folds of the larynx; ? the vocal ligament;

(f) in plants: a sap vessel; also, a vascular bundle of a leaf.

2. A channel of water, stream, rivulet; a spring; a well; an underground water-course; also fig.: a source (of sth.); also, ? a bloodline, one's lineage.
3. A linear deposit of minerals, metal, ore etc., a lode; – of (the) erthe, a substance found in or extracted from a deposit; – of mountaine, ? a cupreous mineral; ? mountain blue, copper carbonate; a (the) – of, a mineral or metal found in streaks or veins within (a stone, a nugget of ore etc.); also, a seam or filament of contrasting texture or color in a stone or nugget of ore; also fig.: ? a natural feature of personality, a propensity [quot. a1425 (?a1400) RRose].

Figure One


aorta
pulmonary artery
sup. v. cava
pulmonary veins
left auricle
r. auricle
left ventricle
r. ventricle
inf. v. cava

The Lexicographer's Dilemma: Handling Mistranslations, Misconstruings, and Paronomasia in Source Texts

Elizabeth S. Girsch

I had been an editor at the *Middle English Dictionary* for only a short while when I encountered the citation that got me started thinking about the questions I address in this paper.[1] I was working away at the Middle English verb *treden*, which means, as one might expect on the basis of its modern reflex, 'to tread', as well as 'to tread upon and thus to despoil', 'to follow a course by walking', and a few other related concepts. It also means 'to copulate with' – hence Chaucer's Host's rueful observation that "religioun hath take up al the corn of tredyng"[2] and his remark to the Monk, "thou woldest han ben a tredefowel

[1] I owe thanks for suggestions and comments to the original audience of this paper, which was delivered at a lexicographical session of the first International Medieval Conference at Leeds in July 1994, as well as to numerous staff members, past and present, of the *Middle English Dictionary* (hereafter referred to as *MED*); I would like to express my gratitude to all as one. It is to be assumed that the Latin cited in all instances is Medieval rather than Classical Latin, but this is of no significance in the arguments and thus has not been specially noted.

[2] (c1375) Chaucer *CT. Mk.*B. 1954-55. For ease of reference to the relevant entry in the *MED*, I have given all references in the form in which they appear in the *MED*. For a complete list of title stencils, along with the bibliographical information on the editions cited, please see Middle English Dictionary, *Plan and Bibliography*. Ann Arbor: University of Michigan Press, 1954, and *Plan and Bibliography, Supplement I*, 1984.

aright",[3] which, while clearly admiring of the Monk's overt masculinity, does hint at a salient feature of this sense: it generally has reference to birds.

The case that I came upon was a line in the Middle English translation of Bartholomaeus Anglicus' encyclopaedic *De proprietatibus rerum* dealing with the stork, of whom Bartholomaeus had observed, according to the translator John of Trevisa, "The male tredeth nou3t the female but in the nest".[4] This renders the Latin "in nido mas feminam non calcat",[5] which, if true, is equally noteworthy if not quite so picturesque a detail as the former. It would appear that Trevisa fell into the admittedly tempting trap of assuming that Latin *calco*, the senses of which neatly overlap those of Middle English *treden* 'to tread or tread upon' and 'to despoil by treading on', also shared with it the sense 'to have sexual congress with' – or maybe he was just influenced by those long-standing associations between storks and sex.[6] At any rate, Trevisa's English does not quite say what the Latin source meant – or what it said – and so, as a relatively inexperienced editor, I went rummaging through the most recently published volumes in search of the received manner of styling a gloss to indicate such a situation.

What I found was a formula reading "mistransl. L --". At first glance that seemed satisfactory, given that there was clearly a mistranslation at hand, but on closer examination I realised that the formula was intended not to substitute for but to accompany a gloss. Furthermore, the citation had to be placed within the framework of the whole definition, and placement does, after all, make a semantic statement itself. So, in this case, was I to place the occurrence under the gloss 'to copulate with', even though in fact I could hardly claim absolute certainty that Trevisa had intended it that way? Or, employing another formula we resort to, ought I to have glossed it as "to tread upon, with punning reference to sense (b)"? The problem with that

[3] (c1375) Chaucer *CT.Mk.B.* 1945.

[4] (1398) Trev. *Barth.*(Add: Seymour) 619/27. For the reader's convenience, references will be to this edition, but in fact the text is taken from a facsimile of MS. BM Additional 27944, in possession of the *MED*. No significant variation exists between the printed edition and the manuscript in any citation I present.

[5] For Trevisa's Latin source, the *MED* uses a photostat of the printed edition of 1483.

[6] The *MED*'s preferred authority for Latin is the *Oxford Latin Dictionary*, from which this paraphrased definition is derived.

solution is that it was not punning, except perhaps in the most Freudian sense. Furthermore, neither formula took account of the fact that Trevisa had reconstrued the entire notion to say "not ... except in" where his source had a simple negation. The upshot of all this indecisive casting and recasting of the entry was that I began to formulate a series of questions that in the end seemed to boil down to one central question: when we are dealing with occurrences of words in translated works in Middle English, do we gloss what they say, what we think that they should say, or what we think the translator meant, whether he should have meant it or not?

Addressing this question to an audience of lexicographers, I might expect a politely couched but covertly disdainful rejoinder: what sort of question is this? Of course the lexicographer's obligation is to gloss what is in the text. Anything else is interpretation, to put it charitably, or editorial meddling, to put it more harshly. My initial impulse, at the beginning of my lexicographical career, was certainly consistent with that principle. After a great deal more experience in the ensuing years of grappling with some pretty intractable problems, however, I have come to the conclusion that this notion of "glossing what is in the text" is not necessarily as straightforward – or even as unquestionably desirable – as it appears, at least when one is dealing with translations. At times, what is truly most important is not what appears in the text itself but what has intervened to produce it, and to the extent that the lexicographer is able to present that, the user of the dictionary will perhaps profit, or at least not be beguiled by something that seems more straightforward than it is.

For a first illustration, we might take up the case of Aethelraed Unraed. As everyone no doubt knows, the Anglo-Saxon wits who first applied this name to the hapless younger son of King Edgar referred not to his youthfulness and inexperience on assuming the throne, but rather to his imperviousness to good advice from his elders, and the point of the sobriquet was its contrast to his given name, i.e., "noble counsel – no counsel". When Middle English historians who commented on his inglorious reign found his name and nickname in their sources, some merely reproduced what they found, untranslated and thus probably lost to their audience as a political comment.[7] But a few dutifully attempted

[7] For example, note c1210 *Leges Edw.Conf.* in Liebermann *Gesetze* 1. 662, which mentions "Etheldredus Unrad", and a reference in the somewhat later *Livere de Reis de Britannie e le Livere de Reis de Engleterre* to "Eilred Unred". For further

to render it into their own language, and there arose an interesting set of difficulties. The name Aethelraed was already opaque anyway, the element *ethel* having dropped from the language, and the sorts of phonological changes that affect opaque compounds ere producing forms like Eilred and Aldredus.[8] But more significant, *unraed* appears to have developed as *unrede*. This noun, while it retained a sense 'bad advice, ill-conceived plan', and so on, was already moribund by the end of the early Middle English period. Meanwhile, there comes into existence a semantically related adjective, *unredi*, with which of course the noun *unrede* might easily be occasionally confused, graphically and perhaps also phonetically. Never common and slated for early loss, this adjective was in conflict with another adjective, the homonymic *unredi*, the most prominent sense of which was 'unprepared' or, as we would say, 'unready'. Now, the larger problem is this: given that right up through the end of the thirteenth century personal names like "Walter Unred" and "Simon Unred" were being recorded, does it appear that the Middle English writers who referred to "Eilred the Unredi" were, in the time-honoured fashion of high-school students in our own era, substituting for *unred*, a term which must have seemed archaic and perhaps unfamiliar – but which obviously had meaning, unlike a mere surname – a familiar adjective, *unredi*? This the lexicographer of course cannot know and thus cannot use as a basis for his treatment of the word, but in a sense he cannot avoid the question, for he has to deal with a practical problem: ought not the forms *unred*, *unrede* and *unredi,* when they render the sobriquet *unraed,* be kept together – but if so, where? Should they be collected under the noun, where they belong historically? Or should they go under the adjective, where the majority of occurrences belong on the basis of form? And if they are placed under the adjective, should they be treated as an otherwise unattested sense 'no-counsel', or should they be referred to the more common adjective which in fact took over, albeit erroneously, in this very usage? As it happened, the sources in which the most problematical forms appeared are not quoted by the *MED*, so the question became moot, but its implications remained troubling to me as I pursued my lexicographical endeavours.

information on the development of the sobriquet, see Keynes in *Ethelred the Unready*, ed. David Hill, BAR ser. 59 (1978), 227- 254.

[8] By the turn of the fourteenth century the medial fricative of the first syllable seems to have been mostly lost, though Latinised forms such as *Etheldredus* also appear.

Not long afterwards I encountered the problem of the hidden pun. The *MED* does, like other dictionaries, have ways of indicating puns in Middle English, when we catch them, as for example in some famous lines from Langland, the *Gawain* poet, Chaucer, and other authors whose works have received minute scrutiny and exhaustive commentary.[9] But when the pun is in the source text and the Middle English translator appears not to catch it, or chooses to translate the less appropriate of two possible meanings, we have no consistent means of communicating what lies beneath our text. For example, when the Aelfrician usage that underlies much of the collection of sermons that the *MED* calls *Bod.Hom.* regularly uses biforms of the present participle of the verb meaning 'to live' – *lybbende* when it modifies an ordinary mortal being and *lifigende* when it appears in association with the Father, the Son, or the Holy Ghost – how do we gloss the phrase *lifiende god* in the Middle English version?[10] Does it contain the same punning allusion to the participle for the verb 'to believe' as the Old English does? Given that Middle English loses the biform *lybbende* early and that *liven* has come to share few forms with *leven,* would Aelfric's subtle choice of words have had the same effect on an Middle English reader and translator as it had had, one presumes, on an Old English listener? If so, might there be some way to communicate it in the treatment of Middle English *lifigende*? Is a failure to do so in some sense a failure to gloss the word adequately in its context? Such troubling questions continued to accumulate, even as new lexicographical problems, such as the glossing of a pun, presented themselves.

Puns are indeed lexicographically troublesome, not least because a modern reader's grasp of the language may not always be sufficient to guarantee that he will recognise them on sight: just as it may be that translators sometimes failed to recognise that they were dealing with paronomasia, sometimes we may well wonder whether we ourselves are. In uncertain instances, it seems to me, the larger context has to control the interpretation, as for example in following case, a macaronic entry in a wildly diverse collection of bits and pieces, primarily glosses, contained in MS. BM. Additional 37075.[11] The glossary entry in

[9] For an example of the *MED*'s handling of these sorts of cases, see *tredinge* ger., sense 3.(b); the formula "--used with punning ref. to sense 5" is quite typical.

[10] c1175(?OE) *Bod.Hom.* 86/2.

[11] (a1500) *Add. 37075 Gloss.*

question reads *de malo in penis: the lenger the wors*.[12] Now, there's a perfectly plausible straight and sober interpretation available for this passage, but it is almost irresistibly tempting to consider reading it, as the editors suggest, as "concerning pain in the penis, the longer the worse". One might say that it is not the business of the serious lexicographer to go digging for off-colour interpretations, but in this case the entry occurs in a manuscript filled with personal notes, seemingly off-hand scribbles, and what may be scribal commentary on the work in which the glossator is engaged: for example, just a few folios further down, completely out of alphabetic or logical sequence, we find *my fote slepyth,* followed (rather than preceded, as is usual) by the perhaps hastily added and presumably legitimising Latin translation, *pes meus somdolet*. There is, furthermore, an exceedingly suggestive macaronic verse on an earlier folio and a Latin verse playing on the two senses of Latin *anus* elsewhere.[13] If there were ever a context in which a risqué pun might be expected, this is it. The legitimacy of the interpretation conceded, however, the lexicographer is still likely to be at a loss as to how to present it: under *longe*? Under *wers*? Under *penis?* In fact the reader will find this quotation under none of the three, perhaps for reasons that are not immediately apparent. An editor could have handled it, perhaps, under *longe* or *wers* as a phrase and used the formula "with punning reference to" – but it is in neither of those words that the pun really resides. Yet had he placed it under *penis*, the *MED* might properly have been accused of including a Latin word of manifestly unrelated sense under our entry merely through what might have been a clang association. But if it is the glossator's clang association, as I genuinely think it may be, rather than ours, should the dictionary not try to give the reader a sense of what we think our native speaker actually meant?

Whatever the decision on this line, at least one can assume that this probable punster was reasonably bilingual in English and Latin. But perhaps not everyone in his scriptorium was. An occasional work survives as linguistic evidence sufficient to dispel any notion that everyone in the medieval community of scholars actually knew Latin. Witness the efforts of the anonymous perpetrator of the Middle English *Letter of Aristotle to Alexander*, a work translated, if one may use the

[12] (a1500) *Add. 37075 Gloss* 68/305a.

[13] *Ibid.* 47/244a and 15/100b.

term fairly loosely, from workmanlike but reasonably coherent Latin into a fantastic collection of disjointed Middle English sentence-like units.[14] Take the following as an illustration:

> Herying thies thynges with a glad soule: "Say yee,
> sumwhat and ever to me
> of mannes lawe mylking theym, what this might be
> that to me so nobly and so
> shynenly and so gretly ye shewen".[15]

This is for Latin

> Tum ego eo responso laetus, "Dicite mihi", inquam,
> "duo senes humanitatis"
> [His auditis animo exhiliratus 'dicite mihi', inquam
> humanitatis] lege eos mulcens
> "quid sit illud quod mihi tam illustre et tam
> magnificum pollicemini",

(as difficult as it may be to see the correspondence). Suffice it to say that the translator seems to have understood *umquam* for *inquam, mulgens* for *mulcens,* and, one can only presume, *publicamini* for *pollicemini.* For such situations, one *MED* editor has proposed, only half-jokingly, a field-label: "in *Let. Alex*". Only a little less jestingly I would suggest a field label-cum-warning: "in mindlessly literal translation". This might hint at the nature of the problem, which is not just that individual words are mistranslated; their sense is entirely lost because the grammatical information that they convey, such as the person of the verbs, is omitted and, even more, because syntactic relations among them are ignored. At the *MED* we make use of a legitimate explanatory rubric, "misconstr. [misconstruing] L --", but it is obviously best suited for the glossing of discrete phrasal units; it cannot capture the effect produced by the omissions and disjunctures displayed in the passage quoted above. The closest parallel I know is the probably-apocryphal re-translation said to have occurred when a machine-produced Chinese translation of the English throwaway utterance "out of sight, out of mind" was run back

[14] (a1500) *Let.Alex.*

[15] (a1500) *Let.Alex.*

through the Chinese-to-English program: it emerged, the story goes, as "invisible idiot".

To reformulate once again the question that appears to be most significant in such cases: what did the translator who produced passages that seem to us nonsensical himself think that he was writing? Did his words make sense to him and, if so, should the lexicographer not assume that they could make sense to him also, if the matter were approached properly? If not, does the dictionary not genuinely have an obligation to indicate, by some means, that the Middle English language is not actually being represented in such hopelessly confused and garbled passages? And what might such means be? A consistent policy of inclusion, such as the *MED*'s, necessitates some conventions to deal with complications involving mistranslations and misconstruings, but these are, as should be clear, not always equal to communicating the actual nature of the difficulty.

Anyone who can comfortably dismiss these questions in relation to the material presented above might want to give the next case a try. It is from a mid-to-late fifteenth century translation of Higden's *Polichronicon*, anonymous, and inarguably an astoundingly ambitious undertaking for someone whose grasp of Latin was not exactly unshakeable. Rendering the straightforward Latin passage

> Ennius poeta ... habitavit in monte Avetino parco
> sumptu et unius ancillae
> ministerio contentus,[16]

Cornishman John of Trevisa, who had himself translated the *Polichronicon* earlier, in the late 1380s, managed to arrive at an equally straightforward piece of Middle English prose:

> Ennius the poete wonede in the hille Aventinus with
> litel cost, and hilde hym
> apayed with the service of oon wenche.[17]

Anonymous, however, produced this gem:

[16] (a1387) Trev. *Higd.* 4.48.
[17] *Ibid.* 4.49.

> Ennius the poete dwelled in the mownte Aventyne,
> was of litelle meite contente
> with the ministery of oon goos.[18]

An editor faced with this passage might reluctantly concede that *ancillae* does look as though it could be a feminine diminutive of *anser*, maybe – but did the translator really think that Ennius kept one as a servant? This is indeed difficult to countenance, even if St Cuthbert was said by his Middle English biographer to have been attended by otters. A closer look at the passage suggests, however, that something else may have been in the translator's mind: where Trevisa took *parco sumptu*, reasonably enough, as 'with litel cost', Anonymous took it as 'of litelle meite contente' and so on – well, one goose certainly equals little meat. Can one make any sense of this otherwise? Maybe it is a howler, but perhaps it is a very strained – but marginally legitimate – interpretation. How is the lexicographer to choose, and, more important, how can one communicate what the choices were? Obviously the quotation must be placed under *gos* (though of course it was not), but is it a "mistransl", or a complete reinterpretation, more or less internally consistent, of the source text?

Similar situations abound in the early fourteenth-century *Ayenbite of Inwit,* Dan Michel's Kentish translation of the French *Le Somme des Vices et des Vertues*.[19] Some instances are readily dismissible, but some require a more complicated analysis. In one case, for example, a misreading begets a mistranslation, or perhaps vice versa: where the translator has

> This zinne of prede ys to dreduol, uor he ablent men zuo thet hi ham zelue ne knaweth ne ne zyeth. Thet is the wel strang and the wel special ald to the dyeule huer-of he begyleth the he3e men and uayre[20]

the French text speaks of pride as the *uins* or 'wines' of the devil which *eniure* or 'intoxicate' men. Having misread *uins* as *uius* 'old', he was compelled to find another interpretation for *eniure,* which he apparently

[18] *Ibid.* 4.49.
[19] (c1340) *Ayenb.*
[20] (c1340) *Ayenb.* 16/25ff.

decided to take as a form of *engignier* 'to deceive'. Well, we have to assume that it made sense to him, but what are we to do with *ald*? What did he think it meant, and was it for him an adjective, as the French word that he thought he was translating was, even though he modifies it as though it were a noun? Should it be glossed under the adjective regardless? There are dozens of such cases; and, while the formula "mistransl. F --" does hint at the nature of the problem, it doesn't really explain what makes such things interesting as linguistic phenomena.

Perhaps that is what is really at the heart of this, for my part, anyway. In my entries I want to explain, to discuss, to clarify, not to mention providing the reader with the fruits of my labours in figuring out what produced the mess that I am glossing so glibly. Only such an impulse could justify comment in a case like this final one, which suggests a possible palaeographic error leading to an easily-corrected false reading – that appears in all the texts and makes perfectly good sense: in the Latin text underlying the early thirteenth-century allegorical homily *Sawles Warde,* the phrase *fur principalis* occurs in a passage representing the Christian's soul as a household to be defended against a band of intruders. In the Middle English texts, Prudence is exhorted to guard against the *head therof.*[21] The overpowering temptation is to say to oneself, "I cannot believe that no one saw that this should have read *head theof*", and then somewhat smugly to supply it in the entry. But that is not what is there in the Middle English text. It should be, I think it safe to say, and surely one can understand the impulse of the editor who, having been clever enough to have noticed it, longs to make sure that the readers see it too. Well, do we owe that to the readers? Or to ourselves? After all, what is there makes sense. What business do we have correcting what is not, strictly speaking, an error? And if we do permit ourselves this kind of interference, how do we signal it? Should *therof* have been glossed under *thefe* and recorded as an erroneous spelling? These considerations led me back to formulating questions, the questions that I think all lexicographers ultimately have to ask in relation to the kinds of problems discussed in this article.

First, we have to ask what kind of dictionary we are producing. The *MED*, though we lapse into encyclopaedic style on occasion and make forays into the realm of descriptive syntax when I am not myself restrained by the editor-in-chief, is essentially a glossing dictionary.

[21] a1225(?c1200) *SWard* 6/36.

Although nowhere stated in plain words, the *MED*'s primary obligation to its readers is understood to be the provision, for each and every citation, of a word or phrase that can serve as an accurate and idiomatic rendering in modern English.[22] It may sound less ambitious than, say, the encyclopaedic approach, but it is actually no small task. And the argument that underlies it is that this is what best serves the typical user, who wants to be able to plug in an English word or phrase that makes sense in the context. But does it? As I look back at my own examples, I am forced to conclude that it may not in every case. After all, *gos* is 'goose', but what does that really tell anyone about the meaning of the passage in question, even if we add "mistransl. L *ancillae*"? If we – and our readers – are interested in the way the Middle English language behaved when in actual use by its actual speakers, we may genuinely need in some instances to provide not just glosses but also rationales for glosses. And perhaps sometimes we may need to provide not a gloss at all but rather an explanation of why we cannot give one.

This could easily turn into a prescription for, or defence of, editorial arrogance, and I do not intend that at all. The trick, it seems to me, is in striking a balance between giving the translators and the scribes credit for greater facility with their own language than we can claim, and acknowledging the possibility that they may occasionally have made extraordinarily stupid blunders. We should always begin by presuming that the person who wrote the words that we are glossing knew what he or she meant to say – and try to figure out what that might have been – but when we have reason to believe that we have come upon an instance when he did not, we ought to say so; that tells the reader something important too. In either case we ought, though, if at all possible, to make our thinking transparent, so that if the joke is really on the editor after all someone can gleefully point it out to us later on.

What I am advocating, one might charge, is less a dictionary than a collection of miniature word-studies, more or less detailed according to the complexities associated with any particular word. Inasmuch as the *MED* in its present terse style will run to more than 16 volumes, a more

[22] The introduction to the *Plan and Bibliography* puts it succinctly: "A MnE translation or paraphrase is often the most effective way of putting the meaning across, and it is by all odds the briefest. Hence this method is used unless the lack of an unambiguous MnE equivalent requires explicit definition of the ME term". In practice, this has come to mean that the ideal entry consists of the smallest possible set of synonyms or synonymous phrases that will accommodate all the citations idiomatically.

expansive approach obviously could never have been practical. Had we gone at this problem systematically, we could at least have expanded our repertoire of rubrics beyond the present set of formulae, but perhaps even that would never have been very realistic, since the more the formulae proliferate the more frustrating and even misleading it becomes when cases continue to elude them. At the *MED*, in any case, we are editing words at the very end of the alphabet; it is too late for us to change now even if change were feasible on all the other grounds. But I am concluding on an optimistic note nonetheless. The world of electronic lexicography is just opening up.[23] Those who become involved in these early stages can profit from our experience and exploit the infinitely more flexible format of the electronic dictionary to give the users a view of the process as well as the product of lexicography. It will be possible, I think, to make the relevant passages of source texts available at the touch of a button. It will be possible to reference an expanded "definition" or to call up additional uses of troublesome words, other phrases containing the word in question, and other useful pieces of information not practical to provide in the entry itself. The user will thus be able to reconstruct what led the editor to his final gloss. Will most readers still just check to see what *viritoot* means in line 3770 of *The Miller's Tale*? Maybe. But a few may follow along the path that concludes in such unrevealing phrases as "mistransl. F --", "misconstr. L --", or "with punning reference to --" and thus discover, along with us harmless drudges, that the lexicographical journey is, as journeys so often are, vastly more entertaining than the definitive destination.

[23] As of this writing, an electronic version of the *MED* (*EMED*) is in preparation under the auspices of the University of Michigan Press.

Lexicography and Philology

Karina van Dalen-Oskam and Katrien Depuydt

INTRODUCTION

A lexicographer normally investigates a set of miscellaneous texts. A philologist usually concentrates on only one – preferably literary – work. Whereas the lexicographer's goal is a linguistic description and a semantic interpretation of all words in his material, the philologist is working towards full comprehension of "his" text, and usually tries to reconstruct its autograph.

The basic assumption of the lexicographer is: the text as such must be correct. Not so for the philologist: he always wonders whether the text is correct. So the lexicographer assumes that the writer or copyist wanted to put down linguistically correct statements, whereas the philologist tries to find out whether the copyist has intentionally or unintentionally changed the original text and its meaning.

If we have not yet been able to convince you of the incompatibility of the two, we will try to do so by adding another very important difference arising from the above-mentioned divergent principles: because the lexicographer does not have easy access to the manuscripts himself, he prefers to make use of diplomatic editions, whereas philologists nowadays normally prepare critical ones. This leads to a paradoxical situation. For his edition, the philologist uses dictionaries that are often based on a relatively high number of old and/or unreliable editions. He will also rely upon linguistic handbooks, which are based on these dictionaries and will therefore occasionally contain faulty citations or reconstructions. In this way, it is not unlikely that in new critical editions errors will have crept in. As again these editions will

probably be used as material for new dictionaries, more corrupted words and phrases will crop up.

In 1933 the famous Dutch archivist Willem De Vreese wrote a paper called "Paradox over den grooten nood der Nederlandsche Philologie" (Paradox on the Great Need of Dutch Philology).[1] In it he gave a rather depressing description of the paradoxical situation for the Middle Dutch language. In his view the source material of the *Middelnederlandsch Woordenboek* (Dictionary of Middle Dutch) by Jacob Verdam and Eelco Verwijs (published between 1885 and 1929) is approximately 80 per cent unreliable (p. 169). Therefore, all kinds of mistakes have caused a considerable amount of misconceptions concerning the Middle Dutch language. He therefore urgently proposed to re-edit the majority of the already published texts or at least collate as many of the old editions as possible, there-to following very strict and uniform rules (p. 175). Unfortunately, the *Corpus Collationum* did not get airborne. More than thirty years later professor Ria Jansen-Sieben looked back at De Vreese's paper in her inaugural lecture "Is De Vreeses paradox nog actueel?" (Does the paradox of De Vreese still exist?).[2] She had to conclude that not much had changed, since the editors stayed with the old critical principles and refused to make diplomatic editions, thereby, as Jansen-Sieben puts it, only corrupting the existing material even more.

In this paper we want to take a look at the current situation in Belgium and in the Netherlands. Is the lexicographer of the computerised nineties still haunted by this paradox? First we will give a short description of the dictionary we are compiling. Then we will try to give a survey of the most frequent errors we encountered in the editions we use for our dictionary. Finally, we will describe the procedure we follow when dealing with corrupt material, while elaborating on the differences between the lexicographical approach and that of the traditional philologist.

[1] W. de Vreese, "Paradox over den grooten Nood der Nederlandsche Philologie". (1933). In: W. de Vreese, *Over handschriften en handschriftenkunde*. Tien codicologische studiën bijeengebracht, ingeleid en toegelicht door P.J.H. Vermeeren (Zwolle, 1962), pp. 142-178.

[2] R. Jansen-Sieben, "Is De Vreeses paradox nog actueel?" In: *Tijdschrift van de Vrije Universiteit te Brussel* 12 (1969-1970), pp. 1-14.

SOURCES

At the Institute for Dutch Lexicology in Leiden we are compiling the *Vroegmiddelnederlands Woordenboek* (Dictionary of Early Middle Dutch), together with two other colleagues. We have had ten years (1988-1998) to complete the dictionary, and we have developed a fully computerised system to do the work with. The source material of the *Dictionary of Early Middle Dutch* consists of all extant Dutch manuscripts written in the thirteenth century. This means that charters as well as literary texts and lexicographic material are included. Thirteenth-century works that have come down to us in manuscripts from a later date are excluded from the description.

Most of the thirteenth-century material has been edited diplomatically by Maurits Gysseling, in his *Corpus van Middelnederlandse teksten (tot en met het jaar 1300)* (Corpus of Middle Dutch Texts (up to and including 1300)), published between 1977 and 1987. He did not include some larger texts which were already available in reliable editions: a register of lent properties from Limburg (*Het oudste goederenregister van Oudenbiezen*), a gospel harmony (the *Liège Diatessaron*) and the only lexicographic work in our material, a Latin-Middle Dutch wordlist from Limburg (the *Glossarium Bernense*).[3] The first two are available in a critical edition, whereas the wordlist is transliterated diplomatically.

PROBLEMS WHEN WORKING WITH EDITIONS

The standard of present-day editions has improved considerably. The diplomatic editions by Maurits Gysseling are of an extremely high quality and the critical editions we need for our dictionary are also very acceptable to work with. That does not mean that a lexicographer should not be aware of mistakes.

Firstly, there are two kinds of printing errors: the errors where the printer's gremlin left printing codes in the final texts, e.g.:

Rijmb. 225,12: Hi sbQltene en*de* was onblide
'He ? and was unhappy'.

where sbQltene *should be* scaltene:

Hi scaltene en*de* was onblide
'He reprimanded him and was unhappy'.

[3] Cf. for this text T.Schoonheim, "The Vocabulary of the *Glossarium Bernense*. Some aspects of the use of a wordlist as source material for a historical dictionary". In: *Amsterdamer Beiträge zur älteren Germanistik* 36 (1992), pp. 131-139.

or made strings of letters disappear, e.g.:

Rijmb. 9,7: **D**Es .v. daghes v*er*sierde god ons here / wen*de* lucht. met groeter ere
'On the fifth day, God, our Lord, created ? air with great honour'.

where wen*de should be* wat*er* en*de* 'water and':

DEs .v. daghes v*er*sierde god ons here / w<u>at*er* en*de*</u> lucht. met groeter ere
'On the fifth day, God, our Lord, created <u>water</u> and air with great honour'.

These mistakes are very easily discovered and repaired. Typing errors however are much more difficult to track down, e.g.:

Rijmb. 308,8-9: Ja ne weetstu wat dese doe goet. / Hi ne p*r*ofitiserd mi nemm*er* goet.
'Indeed, don't you know what this person does right ? He never gives me any positive prediction'.

where doe goet 'does right' *should be* doet 'does':

Ja ne weetstu wat dese <u>doet</u>. / Hi ne p*r*ofitiserd mi nemm*er* goet.
'Indeed, don't you know what this person <u>does</u>: he never gives me any positive prediction'.

A more serious problem is the editors' reading errors. Until now, we have only discovered one error in Maurits Gysseling's diplomatic editions, namely:

Rijmb. 100,1: Dus wueden si egypten land
'So they rage ?? Egypt'.

where wueden '(they) rage' *should be* roueden '(they) ransacked':

Dus <u>r</u>oueden si egypten land
'So they <u>ransacked</u> Egypt'.

The edition of the *Glossarium Bernense* however has proved to be less than reliable in that particular domain. We are fortunate enough though that the text was published together with the photographs of the manuscript so that frequent checks enable us to discover and of course to correct the mistakes, e.g.:

Bern. 149,16: capparis : en bockein
'caper : a billy goat?'

where bockein 'billy goat' *should be* boe*m*kein 'small tree':

capparis : en bo<u>e*m*</u>kein

'caper : a small tree'

Then there are the editors' misinterpretations. Sometimes the abbreviation marks used in a manuscript are not interpreted correctly, e.g.:

Rijmb. 92,38: ia*m*nes *should be* ia*n*nes 'Jannes, a magician and opponent of Mozes (2 Tim 3:8)'.

In critical editions, capital letters can be found in the wrong places, e.g.:

Oudenb. 24,8 and 16: anden Herwech 'along the Roman road'.

The use of a capital letter in the critical edition strongly suggests that this 'Herwech' should be interpreted as the name of a certain road, which is not all together certain, considering the manuscript, that has 'herwech' without a capital letter, implying that both possibilities are in order. Maybe more serious even are unjustly made suggestions for emendation, e.g.:

Nat.Bl.D 242,30 : zjfius

This word form should not be altered to 'xifius' as it is a legitimate spelling variant for Medieval Latin 'xifius' < Greek/Latin 'xiphias' 'swordfish'; e.g.:

Bern. 235,6 jnuoluere : in wolken, in welken

The editors propose as alternative readings 'in winden' respectively 'in wenden' whereas we would simply suggest the possible connection the two word-forms have with Middle High German 'welgen, walgen' 'to roll' which is exactly the meaning of Medieval Latin 'involvere'. The Limburg dialect in which the glossary is written displays a lot of German influence.

There are of course the emendations that are entirely correct from the philologist's point of view but absolutely unnecessary as far as the lexicographer is concerned (we will explore that subject matter a little further on). A good example of that is Gysseling's method of working when editing literary texts. He does not systematically give all variants but in some cases suggests an alternative reading, usually based on one of the existing variants, not mentioning however which manuscript the alternative reading is based on.

It is clear that a lexicographer has to continue to be cautious when using text editions. It is only through working with the material that after a certain period of time he can ascertain the degree of reliability of the editions used. It will however occasionally remain necessary to consult the manuscripts themselves.

TO EMEND OR NOT TO EMEND?

Mistakes have to be corrected. Sometimes a lexicographer also has to suggest emendations or alternative readings. However, the point of view of a lexicographer differs greatly from that of the traditional philologist and he judges material from official documents differently from material found in literary texts. To go into the latter first: official documents that have come down to us are usually autographs or first copies so that it is highly unlikely that these texts will contain many mistakes. The official documents in our material are the first to be written in the vernacular and as they were meant for local use only the language is still highly dialectal. Literary texts had to function more regionally so that the language was more standardised. Therefore, obscure passages (c.q words) in official documents are treated with the utmost caution: it is always possible that the lexicographer is confronted with a dialectal variant he does not recognise as knowledge of that particular dialect or region might be insufficient, e.g.:

Corp.I 1080,51: vergolden den here van totenate
'paid to the lord of ?'

We do not suggest any kind of emendation, hoping one of our readers might eventually identify this place name for us.

In his *Eloge de la variante. Histoire critique de la philologie*[4] Bernard Cerquiglini criticises the traditional philologist who considers every variant to be a mistake since the aim is to reconstruct the autograph of a text. In his essay, he basically points out that variance is the most important characteristic of medieval writing and that it is the degree of variation which editors should be focused on. Each manuscript has a value of its own and displays the joy a medieval scribe had in writing in the vernacular. As lexicographers we have a similar attitude towards our material in that we always begin our analysis of a word-form with the assumption that it is correct. Consequently, our attitude towards emendations will in some ways differ from that of a regular philologist.

Both philologist and lexicographer will correct obvious writing errors, e.g.:

[4] B. Cerquiglini, *Eloge de la variante. Histoire critique de la philologie*. Paris, 1989. For a critical reaction cf. K. Busby, "Variance and the Politics of Textual Criticism". In: *Towards a Synthesis? Essays on the New Philology*. Edited by K. Busby. (Amsterdam/Atlanta, 1993), pp. 29-45.

Corp.I 632,32: gheseghet met onsen seghelen
'said? with our seals'
where gheseghet (said) *should be* gheseghelt 'sealed'.
gheseghelt met onsen seghelen
'sealed with our seals'

In our dictionary we clearly indicate what was found in the manuscript and what should have been written down. In particular mistakes in well-known proper names can be corrected with a fair amount of certainty, e.g.:

Rijmb. 34,41 Mecodius *should be* Metodius 'Methodius of Olympus, a Christian author who died ca. 311'.

Sometimes, for a perfectly normal word or passage, the philologist will suggest an alternative reading based on knowledge of the context and the history of the text. He has every right to do so. The lexicographer however only describes what is actually found in the text and not what probably should have been written down, e.g.:

Nat.Bl.D 394,13-18: Iris die gheliket kerstale (..). / latemenne dorscinen vander zonnen / jn een hus (..). / hi sal maken thant / enen reghenbode ande want
'Iris resembles a crystal (..)., whenever in a house (..). one lets the sun shine through the stone, it will create a 'messenger of the rain' on the wall.

Other manuscripts have the variant 'reghenboghe' 'rainbow'. As a rainbow could be considered as 'messenger of the rain' and Iris is the messenger of the Gods, we chose not to emend but to consider this 'messenger of the rain' to be a very poetical variant. At the very most the lexicographer will make a remark that perhaps something else was originally intended.[5] Even whole passages a philologist might consider to be corrupt remain uncorrected if the passage is grammatically correct, e.g.:

Rijmb. 294,9-10: Die coninc sprac te sinen wiue. Vercleed v kinder
'The king said to his wife: change your children's clothes'.

[5] W.J.J. Pijnenburg, "Lexicografie en de studie van het Middelnederlands". In: *De studie van de Middelnederlandse letterkunde: stand en toekomst.* Symposium Antwerpen 22-24 September 1988. Onder redactie van F.P. van Oostrom en F. Willaert. (Hilversum, 1989), pp. 45-54, especially pp. 50-52.

Other manuscripts, including the Holy Bible have something like: verkeert v cleeder 'change your clothes'. Again we chose not to emend, only to point out the difference.

Totally obscure words or passages in a text will be labelled as such by the lexicographer but an alternative reading will not always be suggested. He might quote the parallel passage in the source text, when dealing with a translation, or even give the readings of the other extant manuscripts to indicate what might have been the right content. The lexicographer will only emend when, supported by the variants or the source text, he can propose a reading with the highest degree of certainty, e.g.:

Rein.E 387,13-15: [hoe mochti]c v*er*getten dis / [dat aldar s]ininit (or: [s]imnit ?) die sies / [die valsce p]ennincge sloeh
'How could I forget that over there S...? the hound stroke false coins'.

When consulting the other existing manuscripts of this text we learn that '[s]ininit' as well as 'sies' have different variants:

reynout de ries 'Reynout the fool'
symon die sies 'Simon the hound'
Symonet die rike vriess 'Simonet the rich Frisian' and
Pater simonet die vriese 'Father Simonet the Frisian'.

A philologist will probably feel the urge to try and reconstruct what may have been written down originally. From a lexicographical point of view however, only [s]ininit deserves a reconstruction. To that end, we have chosen 'simonet' as the most probable reading for our manuscript, as this form is graphically closest to [s]ininit (or [s]imnit).

CONCLUSION

Summarising: the quality of the lexicographer's work depends on personal knowledge and ability to value the knowledge and skills of the editor of the source material.

The source material of the *Dictionary of Early Middle Dutch* is of much higher quality than that of the *Dictionary of Middle Dutch* a century ago. This implies a greater degree of reliability. However, we have to ask ourselves to what extent we can check and correct the editions we use as source material. Heinz Antony describes in his article "Korruptel oder Lemma?" that his efforts to collate with the manuscripts themselves as many questionable readings in the editions on which the *Mittellateinisches Wörterbuch* is based have prolonged the planned period

in which his dictionary has to be completed considerably.[6] However, to complete the *Dictionary of Early Middle Dutch* we have very limited time. We might even ask ourselves whether it is entirely justified to withhold the new dictionary from the philologist any longer than necessary. The sooner that philologists can make use of the new work of lexicographers, the sooner they will be able to produce better editions which will again function as material for still better dictionaries. We would strongly recommend furnishing *diplomatic* editions, preferably with a list of variant readings. Suggestions for reconstructions should be limited to the footnotes, and *never* be included in the text itself.

We hope that in time more and more philologists will begin to realise that they personally influence the reliability of the dictionaries and handbooks they use.

[6] H. Antony, "Korruptel oder Lemma? Die Problematik der Lexikographie auf dem Hintergrund der Editoren". In: *Mittellateinisches Jahrbuch* 16 (1981), pp. 188-233.